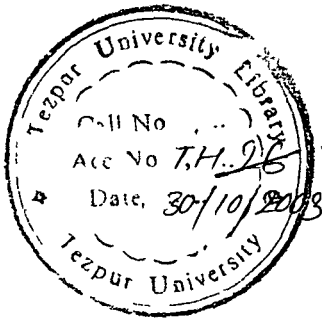
T21

TH-26 43026

ASSESSMENT OF TEA QUALITY : SOME STATISTICAL ANALYSIS

Sanjoy Ketan Paul

A DISSERTATION SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN STATISTICS



TEZPUR UNIVERSITY
TEZPUR , ASSAM
INDIA
2000

Dedicated To My Teacher
Manoranjan Pal

This is to certify that the thesis entitled "**Assessment of Tea Quality : Some Statistical Analysis**" submitted by Sri Sanjoy Ketan Paul, who got his name registered on for the award of Ph. D. (Science) degree of Tezpur University is absolutely based upon his own work under the joint supervision of Dr. Munin Borah of the Department of Mathematical Sciences, Tezpur University and Dr. Manoranjan Pal of Economic Research Unit, Indian Statistical Institute. Neither this thesis nor any part of it has been submitted for any degree/diploma or any other academic award anywhere before.

Signatures of the supervisors with seal

Dr. Munin Borah
The Head
Department of Mathematical Sciences
Tezpur University
Tezpur, India
Date:

Dr. Manoranjan Pal
Associate Professor
Economic Research Unit
Indian Statistical Institute
Calcutta, India
Date:

This is to certify that all the organoleptic and biochemical data analysed in the study entitled "Assessment of Tea Quality: Some Statistical Analysis" by Sri Sanjay Ketan Paul was provided by the Tea Research Association.

He has presented his findings to us in `in-house' discussions and seminars from time to time. The Association is of the view that the findings of Sri Paul, being new to tea industry would generate considerable interest.

Secretary

Secretary,
TEA RESEARCH ASSOCIATION

TEA RESEARCH ASSOCIATION Calcutta

ACKNOWLEDGEMENTS

I cannot but recollect the memory of my beloved father who passed away last year. It is he who encouraged me the most along with my mother and sisters. However, the work would not have been completed without the great support of my teacher Dr. Manoranjan Pal of Indian Statistical Institute. He has not only monitored the whole work from the very beginning but tirelessly helped bit by bit to develop statistical ideas to formulate a chemometric problem methodologically. I am especially indebted to my supervisor Dr. Munin Borah of the Department of Mathematical Sciences, Tezpur University for his constant support in this study.

I am grateful to the Director of Indian Statistical Institute, Calcutta, for providing me all possible research support to conduct this study in the Institute. Thanks also due to the Head of the Department of Economic Research Unit of the Institute.

The whole study is based on the chemical and sensory information on CTC teas provided by the Tea Research Association of India (TRA). I am grateful to the Director, the biochemists and the Tea Taster Mr. Ajit K. Das, in particular, of TRA. I can never forget the great technical support provided by the Secretary of TRA for this study. I am also grateful to the Tata Tea Ltd., India, for providing one CTC tea data for this study. Thanks also due to Dr. Premananda Bharati of the Anthropometry and Human Genetics Unit of Indian Statistical Institute, Calcutta.

I simply cannot escape expressing my gratitude for the enlightenment I got from Professor John C. W. Rayner of University of Wollongong, Australia and Dr. Tormod Naes of MATFORSK, Norway, from the long and fruitful discussion with them. Constructive comments and criticisms made by anonymous referees of some Journals (in particular, Journal of Science of Food and Agriculture, Food Quality and Preference and the Journal of Applied Econometrics), who kindly reviewed some parts of the draft in paper form in course of publication, are also gratefully acknowledged.

Contents

Page

Page

Page

ABSTRACT

Tea is a unique commodity in many respects. Like other agricultural produce, the valuation of tea, though an agricultural product, is not determined by the free play of forces of supply and demand. Primarily, the valuation of tea depends on its quality. The success of tea producers in global tea trade mainly depends on the quality of their product. In the changed global economic scenario, new equations integrating production, quality management and market forces have been evolved. The consumers, especially from the west, are very much concerned about the quality of food and beverage products. That's why it has become extremely necessary for the tea producers to adopt a balanced and reliable approach to the quality management. The quality of tea, like other beverages, is evaluated by sensory methods. The Tea Tasters play an important role in the tea trade by judging the tea lots in terms of overall quality in auction centers. The Tasters of Broker Houses taste the infused sample, and ultimately give the basic price for different lots of teas. Though the judgement of quality by these Tasters matters a lot in the evaluation of the produce, it is not devoid of human bias of various kinds resulting in increasing uncertainty. The Tasters may have some preconceived idea about the origin and chemical composition of the tea, there remains every possibility of human bias which may have significant impact on the price and demand for a particular brand of tea. The tasters consider liquor characteristics, such as strength, brightness, briskness, flavour, etc. resulting from the infused tea, while assessing tea. A group of Tasters may broadly agree on the merits of a sample of tea, but there is bound to be some difference of opinion among themin details. We note here that market conditions in general and the requirements of the Broker House whose needs the Taster serves have profound influence on the judgement of quality. From scientific point of view it is quite reasonable to say that the biochemical parameters inherent in the tea leaf are responsible for the quality in tea. At the present stage of knowledge, most of the biochemical quality parameters are measurable with high degree of accuracy. But we can not quantify the quality on the basis of biochemical information. The Tasters evaluate the different quality attributes (which are due to biochemical parameters) only. That's why, if the correspondence between

biochemical information and the Tasters' judgement can be established, the problem of uncertainty can be reduced. It is important to study the subjectivity of the Taster's choice in assessing the quality of tea. We have discussed different aspects of sensory evaluation and the statistical methodologies to study the error associated with Taster's choice. A sensible approach in this direction is to consider a panel of Tasters, who may assess the same set of tea samples and give scores on each sample independently. We may study the individual variations due to Tasters. A possible model can be thought that would give us an idea about the true inherent quality for the given set of samples after eliminating the bias due to Tasters. It may be possible to measure the extent of relative bias due to individual Tasters also.

From earlier discussions we may legitimately say that there are two aspects in the problem of tea quality assessment. One aspect is the sensory evaluation. The other aspect is the biochemical information. It is believed that the quality attributes are the effects of some combinations of biochemical information. Thus we may think of relating the sensory evaluations with the biochemical quality parameters, and study how the different chemical parameters are related to the Taster's evaluations. If we assume some functional relationship between the sensory evaluations and the biochemical information for a particular set of tea samples, we may possibly write the relationship in the following way:

True Quality = Taster's assessment + Error due to Taster = Quality explained by biochemical parameters + Random error.

We explicitly state here that there may be various ways to approach the problem, we in our study limit the scope to only a few statistical investigations.

Primarily our aim is to study how closely the biochemical information can be functionally related to Taster's choice and thereby study whether the chemical information is being explained by the sensory analysis, may be partially. For this we have associated the biochemical parameters with the choice made by a single Taster. The performance of different regression techniques (e.g. Minimum Absolute Deviation technique, Ridge regression, etc.) is compared with the ordinary least square (OLS) regression. Also the scores on different quality attributes given by a single Taster are associated with the

biochemical information introducing Multiple Response Regression model. The possible non-linearity in the data is studied separately using Box-Cox transformation model. Some critical investigations have been performed to assess the small sample estimation problem with Box-Cox transformed model.

As a distinct approach to study the subjectivity of Taster's choices, we have obtained repeated observations on the quality assessment by taking more than one Taster on the same set of sample. Subjectivity in the sensory evaluation is studied introducing one-way and two-way variance components (VC) models with heteroscedastic formulation of variance-covariance matrix. The next step is to associate the sensory evaluations with the biochemical information after eliminating the bias due to Tasters. This is tried by developing regression models with repeated observations on the response variable. The one-way and two-way error component models have been studied under heteroscedastic environment. The first chapter introduces to the problem under study. Some interesting information about tea is given. The quality aspects in tea are discussed in detail including the biochemical front. Different aspects of sensory evaluation, that is, tea tasting is discussed in detail. This is followed by explanation on the objective and organization of the problem. The next section includes explanations on the different data sets used in this study, which are provided by the Tea Research association of India (TRA) and the Tata Tea Ltd.

A review of the up-to-date literature on the studies made on different aspects of VC models is presented in the second chapter. Special emphasis is given on the repeated measurement models. A detailed account of linear and non-linear models is given. Different aspects of longitudinal data analysis is discussed including approaches to data diagnostics. The ANOVA, MANOVA and generalized MANOVA (GMANOVA) model and estimation of variance components are discussed. This includes a section on profile analysis also. The general linear model for longitudinal data analysis is discussed in detail. The maximum likelihood (ML) and restricted maximum likelihood (REML) method of estimation are discussed along with the available algorithms for estimation. A detailed discussion on the testing aspects in general linear model is given. The review work is done keeping in line with the different repeated measurements models investigated by us to study the quality

aspects of tea.

The third chapter deals with some simple regression studies to associate quality parameters with taster's choice for given set of CTC tea sample. This is attempted using the data set where a single Taster makes sensory/organoleptic evaluations on different samples. Two different problems are dealt with. The first approach is to associate a single quality attribute (say, strength or overall value) with the biochemical information. We have compared the performance of OLS, Ridge regression and the minimum absolute deviation (MAD) method of estimation. The second approach is to associate the sensory evaluations made by a single Taster separately on 'strength', 'quality' and/or 'overall valuation' with biochemical quality parameters. For this we have considered a multiple response regression model. The testing aspect of different parameter estimates is discussed. The 'within - sample' forecasting is also done.

The possible non-linearity in the tea quality assessment data is studied using Box-Cox transformation model in the third chapter. Certain problems of using Box-Cox transformed linear models in case of small samples is discussed in the forth chapter.

The fifth chapter explicitly deals with the different aspects of sensory panel data. The subjectivity of Tasters' choices is studied using repeated measurement variance component models. Measurements made independently by different Tasters on quality attribute constitute the repeated observations on the sensory scores. The ANOVA and ML estimators of the heteroscedastic variance components are obtained.

The one-way random effects linear regression model with repeated observations on the response variable is the theme of sixth chapter. The aim is obviously to associate the biochemical parameters with Tasters' scores. The measurements on biochemical quality parameters are however fixed for each sample. The effects due to Tasters are assumed to be random. With such a formulation we aim to assess the error due to sensory evaluations and also to assess the statistical significance of the effects of different biochemical parameters in quality assessment. The error variances associated with different Taster's choices may be obtained using a heteroscedastic formulation. The ML and REML estimates of regression coefficients and the variance components are obtained employing a simple iterative algorithm. The statistical properties of the regression coefficients are also

discussed.

A generalization of the linear random effects model is done introducing dummy variable to study the variations due to groups. This formulation is done to support a 'quality improvement experiment' for CTC tea, conducted in the Tocklai Experimental Station of TRA. A controlled experiment was carried out to develop a particular brand of CTC tea with better quality. A detailed discussion on the experiment is presented with original data. The experimental samples showed better quality as compared to the control sample.

The one-way random effect model is an appropriate specification if we are drawing samples randomly from the same population of teas. But such a formulation would not be appropriate if we are focusing on a particular set of samples, for example, the samples representing different clones of CTC tea. Inference in this case is conditional on the particular samples under consideration. In this case the effects due to samples (fixed effects) needs to be incorporated in the error component model along with the effects due to Tasters. Thus we have a two-way error component formulation, which is discussed in the seventh chapter. We note here that certain computational problems arise in obtaining the estimates of implicit parameters from the general likelihood function, due to the fact that the design matrix is not of full rank. An alternative ML estimation of variance components and the fixed parameters are discussed.

The last chapter includes brief discussion on the possible statistical studies that could have been undertaken had the required data base on quality assessment been available and some concluding remarks. The data and other technical information for this study are provided by the Tea Research Association of India (TRA).

CHAPTER -1

THE PROBLEM AND BACKGROUND INFORMATION

1.1 Introduction

Tea is a unique commodity in many respects. The valuation of tea, unlike other agricultural products, is not purely determined by the free play of forces of supply and demand. The valuation of tea is considerably distorted by an institution called the **Tea Tasters**. The Tasters play an important role in the tea trade by judging the tea lots about their qualities (or overall qualities) in the auction centers. The tea Broker Houses have their own Tasters who taste the infused tea samples, evaluate the samples and give the basic prices for different lots of teas to be auctioned. Thus the Tasters, who are representatives of the Brokers, reflect the taste and preferences of the ultimate consumers of tea. Though the judgement of quality by these Tasters matters a lot in the valuation of the produce, it is not devoid of human bias of various kinds. The Tasters may also have some pre-conceived idea about the origin and chemical composition of the tea, there remains every possibility of human bias which may have significant impact on the price and demand of a particular brand of tea. Hence there is an urgent need to minimize the uncertainty factors in the quality-price-demand relationship. This can be, to some extent, achieved through an objective and scientific analysis of the influence of the human as well as the chemical factors in tea. From the scientific point of view it is quite reasonable to say that the biochemical parameters inherent in the tea leaf are responsible for the quality in tea. At the present stage of knowledge, most of the biochemical parameters responsible for the quality in tea are measurable with high degree of accuracy. But we can not quantify the quality of tea on the basis of biochemical information, as we do not know the exact relation between quality and the biochemical parameters. There are different quality attributes in tea such as 'strength', 'briskness', 'brightness', etc. that are due to different biochemical parameters. But these quality attributes are judged (evaluated) by the Tea

Tasters only. That is why the correspondence between the biochemical information and the Taster's judgement is called for.

1.2 Some basic information about tea

The history of tea began in ancient China over 5000 years ago. The tea, what we drink today, was discovered by Chinese Emperor Shen-Nung in 2737 BC (ref: web site www.stashtea.com/facts.htm). According to legends, Shen-Nung was boiling a kettle of water in his terrace when the wind blew a few errant tea leaves in to the kettle. The curious Emperor sampled the steaming liquid. He was delighted with its pleasant aroma and taste, and soon the taste of tea spread to Japan and other Far East countries. Early Dutch and English colonists introduced tea to the new world. It was the famed East India Trading Company that formally introduced tea to England and other European countries.

Tea is a beverage made from the processed leaf of a plant whose scientific name is **Camellia Sinensis**. Compared to other agricultural crops, tea production is unique so far as its plantation structure and the processing system are concerned. It takes five years for a tea bush to grow before it reaches any commercial significance. The life of a tea bush is more than fifty years. To maintain productivity and yield, 2% of tea bushes need to be uprooted and re-planted every year. Once the bushes are uprooted, it will be seven years before a re-planted bush will reach commercial bearing. Tea production requires considerable technical expertise and innovation. The Indian tea industry has developed considerable R & D, and has made significant contributions in several areas such as tea biochemistry, biotechnology and agronomy. Specialized research institutes such as Tea Research Association (TRA, with their famous research laboratory at Tocklai, Assam), Darjeeling Tea Research Center, and the Indian Institute of Plantation Management has significantly contributed to the tea research. The laboratory of Tea Research Foundation of Kenya (situated at Kericho), the Tea Research Institute of China, the Shizuoka Tea Experimental Station (Shizuoka, Japan), and the Tea Research Institute of SriLanka (St. Colombus), are some of the famous laboratories in this field. There are more than three

thousand varieties of teas, each with its own flavor, body, color, and aroma. While there is only one species of tea plant, namely Camellia Sinensis, from which all teas are made, local conditions in the various tea growing regions of the world determine varieties, which are unique from each other. The major tea producing countries include India, Sri Lanka, China, Japan and Kenya. We outline below the different types of tea produced in these countries.

Black Tea (Fermented Tea)

Orthodox black tea : Darjeeling tea of India, Keemum tea from China, and Uva tea from Sri Lanka are the most famous black tea in the world because of their superior flavors. There are flavor characteristics that clearly differentiate each from the other. The difference in the aroma characteristics of Keemum, Uva, and the Darjeeling teas are due to the varieties of tea plants used in producing these teas. The processing of this tea is that the harvested leaves are first withered and then rolled, which liberates the aromatic juice and onsets a mysterious chemical change through the absorption of oxygen. This fermentation process occurs in high humidity and warm temperature and turns the leaves a bright copper color and imparts them with subtle flavors.

Crush-Tear-Curl (CTC) tea : The best quality CTC teas are produced in Assam of India and some parts of Sri Lanka and Kenya. The production of CTC tea is rapidly increasing with the increased use of tea bags throughout the world. The flavor of CTC tea is inferior to that of orthodox black tea. CTC teas are stronger. The biochemical aspects of CTC tea are discussed by Yamanishi (1995), and Deb and Ullah (1986), among many others.

Oolong tea : The process is similar to black tea, but the withering and fermentation times are cut down. This type of tea is generally produced in Japan and China. This tea involves the qualities of both black and green tea.

Green tea : The leaves are harvested and immediately put in a large steamer and heated. The leaves are rolled until crisps. They remain green in color. There are some other types of teas such as **white tea, scented black tea**, etc.

The Processing of Tea

The tea leaves (top two leaves and the bud) are first plucked and then brought to the tea manufactory where they undergo the following processes:

The plucked leaves are first withered (dried) on a rack. This withering process may take 10 to 30 hours and its purpose is to bring down the internal moisture of the leaves to somewhere between 60% to 70% of the original moisture. The next phase is the activity of grinding or breaking machines, which cut or crush the leaf. This is done to expose the enzymes present in leafs for further development as a result of coming into contact of oxygen. This is called oxidation. The leaves turn to bright copper penny color and 2 to 3 hours are generally enough to accomplish this. After this phase the tea goes into the drying operation. The tea is dried for between 30 minutes to several hours. This drying operation is very important in that this is the process which 'seals in' all of the flavor and aroma and can determine one of the major differences between a mediocre tea and a high quality tea though they may come from the same plantation. Finally the tea may be exposed to electric roller or other devices to remove the unwanted leaf stem or vein fiber.

The above are the basic steps for producing black tea. There could be several variants of this approach (depending on the production region), but essentially all that is needed are the above steps. We note here that the green and semi-black teas are processed differently than above.

Tea Leaf Grading

Grading tea leaves is very complicated and is done differently in different countries. The most extensive grading is found in black teas, followed by green teas. Black tea is classified into four different categories. The main division is between the leaf grade and the brokens grade. A lesser quality grade of very small pieces is called fannings. Finally dust grades are used primarily in tea bags. In addition to the grading process, the tea industry classifies tea leaves according to the place of plucking, and also the time of year of the harvesting.

1.3 The Quality Aspects in Tea

1.3.1 *Assessment of quality*

The term quality has different connotation for different products. In tea, it is really a complex situation, so far the understanding of quality is concerned. In assessing tea quality, dry leaf appearance is used as one of the criteria. This is because it gives an idea of the standard of the manufacture. But a tea sample is mainly judged from its liquor characteristics. We note here that the appearance of the dry black tea particles varies according to the method of cultivation, manufacturing methods and the skill shown in both. The liquor brewed from the particles or samples varies in taste according to manufacturing method, particle grade, original planting materials and the environmental factors, including the time of year (Baruah, 1992). The work carried out at the Tocklai Experimental Station since early twenties shows that the tea liquor can be described adequately from its liquor characteristics. These are: color, strength, brightness, briskness, flavor and quality. A study of these characteristicsshows that they offer little hope of being translated, at their face value, into quantitative definitions (Harder, 1956; Trick et al., 1967; Baruah, 1992).

Color and brightness have the usual meanings. Lightly fermented tea will tend to be greener, while the liquor of more heavily fermented black tea will be red. The way the tea is fired will have an influence on the liquor's brightness. Lightly fired tea will be comparatively brighter or clearer. Appreciating the beautiful color of tea is intimately connected

to experiencing its wonderful flavor. Flavor is the most important factor in determining the tea quality, and especially the quality of orthodox tea produced in Darjeeling and Dooars regions of India. The market price of orthodox tea is based on its flavor. The term flavor involves both taste (nonvolatile compounds) and aroma (volatile compounds). Aroma is considered to be the most important factor contributing to the quality of tea. Very complicated mixture of volatile compounds, such as terpenoids, alcohols, carbonyl compounds, etc., contributes to the characteristic tea aromas. Yamanishi (1995) makes a detailed study on the flavor of tea.

The degree of briskness has not yet been estimated chemically. It is a sort of astringency. The term 'quality' is rather ambiguous attribute and conveys different meanings to different persons. Weight and Gilchrist (1961) have described quality as a 'dulcet freshness', and used various analogies to describe the term more clearly. In the widest sense, we may say that the quality describes the appeal to the palate as a whole. But in a restricted sense it is described as a liquor characteristic recognizable by a Tea Taster.

The definition of quality may be differing, but it is a fact that the concept of overall quality alters with the kind of tea and its place of cultivation. It appears that the tea trade, which is accustomed to deal with teas from different countries and regions, looks for certain specific character in tea from a particular region. Absence of the regional characters reduces the value of teas.

1.3.2 The Biochemical Quality Parameters in Tea

The characteristic taste of tea is made up of a balanced mixture of astringency, bitterness, bothery taste and slight sweetness. Principal contributors of astringency and bitterness are catechines and caffeine. We note that catechines are phenolic compounds that occur in plants naturally. One of these, polyphenol oxidase, is responsible for turning freshly picked tea leaves black. On the other hand, caffeine is an alkaloid, which is nitrogen containing compound. The alkaloids taste bitter. In the tea brew, part of the caffeine complex with flavanols and play an important role in the tea taste, with contributions to briskness, mouthfeel, and thickness.

Different types of flavanols constitute a group, all of which occur naturally in plants. In the manufacturing process of black tea, some of the catechines are changed to the two biochemical quality parameters, which are said to be the most important quality parameters in tea. These are theaflavins (TF) and thearubigins (TR). Roberts (1950), who originated these terms, considered that the TR are as important to the flavor and quality of tea as are TF. TR is responsible for body, richness, and fullness of the tea brew. TF imparts the mouth sensations of briskness, freshness, and aliveness. Recently, a study of Tea Research Institute of China reported on the influence of catechines and TF on the astringent taste of black tea brew. The study included tea samples from Darjeeling, Assam, China and Kenya. TF produces a yellowish or golden color in black teas. TR is actually the red or brown pigment in tea leaves that are responsible for the color of the tea. Brown pigments occur in the Indian teas and the redpigments occur in Chinese teas. Generally, for CTC tea samples, apart from TF and TR, the other biochemical parameters such as Caffeine (C), Water Soluble Solids (WSS), Total Liquor Color (TLC), Ash content, etc., are also measured. We note that all these biochemical quality parameters are measurable with high degree of accuracy using High-Performance Liquid Chromatography (HPLC) machine.

1.4 The Tea Tasting

The professional Tea Tasters in the world tea trade play a great role. The Tea Tasters play an important role in the quality assessment of tea. A Taster tastes the infused teas and describes each sample in terms of 'strength', 'quality', and 'overall value/quality'. The overall quality score is given considering all the tea liquor characters together. The Tasters' evaluation of tea samples is called the **organoleptic** evaluation or the **sensory** evaluation. The Tasters also indicate the cash valuation of tea. The prices of different lots of tea in the auction centers are guided by the organoleptic evaluations made by the Tasters. We note that the 'cash valuation' of tea samples evaluated by the Tasters may not be the eventual selling price.

While assessing the valuation of tea samples, the Tea Tasters take into consideration mainly the liquor characteristics, such as their color, strength, brightness, briskness, flavor, quality, resulting from the infused tea. The Tasters assess tea by their sensory methods: eye (sight), tongue (taste), nose (smell). The quality perceived by the eye and the tongue are collectively called the black tea quality parameters (Owuor, 1995).

1.5 Objective and Organization of the Problem

As discussed in the introductory part, the valuation of tea mainly depends on quality. There are several quality attributes, which collectively determine the overall quality. This overall quality generally guides the auction market. The quality attributes in tea are due to several biochemical parameters. By this we mean that different biochemical quality parameters present in a particular type of tea (produced in a particular region) give some idea about the strength, color, brightness, etc., of the tea. Obviously, these quality attributes are not directly measurable. This is a big problem, and the chemists in different research stations of the world working on tea, have been struggling for a long time with the problem of chemical evaluation of tea quality. The advances made in the field of tea biochemistry since the last world war seems to be no way nearer to this goal.

That is why, a big importance is given to the tea tasting aspects. In the world tea trade, the assessment of quality for different regional teas basically depends on the sensory or organoleptic evaluations made by the professional Tea Tasters. For a given set of tea samples, a panel of Tasters may broadly agree on the merits of the samples, but there is bound to be some difference of opinion among them in details. This is obvious since the Tasters are human beings, and can not remain aloof from extraneous influences. In many experiments, it has been observed that the Tasters' choices vary widely even for the same quality cup of tea. Market conditions in general and the requirements of the particular buyer of blender whose needs a Taster serves have profound influence on the judgement of a taster. Another important aspect is the absence of a fixed standard of reference for the Taster to compare the teas. As we discussed earlier, the quality varies over regions

and also over the storage periods. In fact inthe auction center, the tea lots lose in price if they remain unsold even for 3 to 4 days only. For these reasons, a Taster is compelled to adopt comparative standard.

From the above discussions, it is clear that while assessing value/quality of a tea, apart from the inherent overall quality (natural quality or true quality), the Tasters keep in mind the market demand structure for particular brand of tea. This comes from the consumer's attitudes towards different regional kinds of teas. In fact, the market demand structure has profound influence on the Taster's judgement. Thus we can not expect that the Taster's choice would only reflect the true inherent quality of a particular tea sample. Reliability of Taster's choice may be questioned on these grounds. Even in case of laboratory experiments, intended to study some tea clones of same region or different regions, the Tasters' scores vary significantly from one to another. The possibility of Taster's effort to promote some particular clones (may be of some particular region also) can not be denied.

It is important to study the subjectivity of Taster's choice in assessing the quality of tea. We may search for different methodologies to study the error associated with the Taster's scores. But if for a given set of samples only one Taster evaluates the teas in terms of different quality attributes, we have no option but to opt for this score only. We must have repeated observations or choices made on each sample to study if there is any variation due to Tasters or due to repeated observations. Two different situations may arise. First, a single Taster may assess the quality of a particular tea sample on different occasions. The situation may be that, for a particular sample, different cups are prepared and the same Taster evaluates each of the cups. We note here that if such repeated observations are made on different days, there will be some variations in quality due to storage effects. Again if he assesses the different cups of the same sample on the same date, he may not remain totally immune from the impact of the first cup while assessing the second cup, and so on. Hence, his evaluations of different cups may not be completely independent. Thus, considering the repeated choices made by single taster, we may not be able to assess the true quality given the sample set.

The second approach is that, we may consider a panel of Tasters, who may assess the same set of samples and give scores on each sample independently. This makes sense. Because, in this case we may overcome the problems discussed above. Also we may study the individual variations due to different Tasters. We may think of a possible model that would give us idea about the true inherent quality (or qualities) for the given set of tea samples after eliminating the bias due to Tasters. It may be possible to assess the extent of bias due to individual Taster. Any way, question may arise on the cost effectiveness of considering a panel of Tasters. But if the industry intends to depend on the Taster's choice for auction pricing, they may always make a choice among the Tasters. To rationalize the whole system of quality assessment, experiments may be conducted for a reasonable length of time to study the error associated with different Tasters. The particular Taster with consistently minimum error of assessment may be the ultimate choice for the industry.

From the introductory discussions we may legitimately say that there are two aspects in the problem of tea quality assessment. One aspect is obviously the tea tasting, which is the sensory or organoleptic evaluation of tea quality attributes. The other aspect is the biochemical information. The question is whether we can say something about the true or inherent quality (in overall sense) of tea only on the basis of biochemical information? At the present stage of knowledge it seems impossible. Again, can we say something about the quality of a given set of tea samples on the basis of sensory evaluations? The Tasters' evaluations explicitly involve the consumers' attitude apart from the actual quality of tea. If a panel of Tasters evaluate the samples, then it may give some idea about the actual quality of the given tea, after eliminating the relative bias due to Tasters.

It is clear from section 1.2, that, the quality attributes (evaluated by the Taster) are nothing but the reflections of different biochemical quality parameters. We may say that the quality attributes are the effects of some combinations of biochemical information. Thus we may think of relating the sensory evaluations with the biochemical quality parameters, and study how the different chemical parameters are related to the Taster's evaluations. If we assume some functional relationship between the sensory evaluations and the biochemical information for a particular set of tea samples, we may possibly write

the relationship in the following way:

True Quality = Taster's Assessment + Error due to Taster = Quality Explained by Biochemical Parameters + Random Error.

Let us explicitly state that there may be various ways to approach the problem, we in our study limit the scope to only statistical investigations. There are a number of published attempts to correlate in a quantitative manner, the chemistry of tea with the Taster's descriptions and cash valuations. Attempts have been made by researchers to explain quality and various liquor characteristics of manufactured teas in terms of chemical composition and biochemical behavior of the unprocessed tea shoots and manufactured teas. To name a few: Harrison and Bose (1942), Roberts (1944), Ramaswamy (1963), Wood and Roberts (1964), Bhatia and Ullah (1965), Biswas and Biswas (1971), among many others.

Roberts (1958) found that TF and TR were largely responsible for color and strength, and that TF were factors in quality and briskness. He also found that highest cash values were given to teas with high TF levels, so long as the TR content was also at satisfactorily high level. Wood and Roberts (1964) observed that Taster's scores for color and strength were related to the TF and TR contents of the manufactured teas. They also observed that scores for briskness and quality depend to some extent on TF, with Caffeine contributing to briskness. According to their observation, cash valuation would be more closely related to TF than to TR. Wickremasinghe and Swain (1965) discussed the relationship between the quantities of phenolic compounds and commercial valuation, and the contributions of the volatile compounds to flavor of Ceylon tea. They observed that the quality of black tea might be predicted from an estimation of the polyphenol content before processing the tea shoots because the amount of polyphenols in black tea depends on the amount originally present in the unprocessed tea shoots. All these studies were usually made on the basis of total correlation between the individual biochemical constituents and the Taster's scores on the individual liquor characteristics or on the cash valuation of the manufactured teas. A much known study on the statistical association of liquor characteristics with the cash valuation of N-E India black tea is due to Biswas and

Biswas (1971, I and II). They used multiple regression technique to determine whether the term 'quality' of the N-E Indian plains black tea has its own single characteristic as recognizable by Tasters or if it is the integration of some of the other important liquor characteristics. They tried to determine the influence of different quality characteristics on the cash valuation of tea. According to their observation, the N-E Indian plain black tea was found to depend mainly on the briskness, quality being increased with an increase in briskness. Cash valuations of CTC as well as Orthodox teas, in general, depended mainly on the quality and/or briskness. They related the biochemical quality parameters with individual Taster's choice and studied the significance of different biochemical parameters. We note here that they did not consider the inherent subjectivity of Taster's choice. **We were largely motivated to initiate statistical study on the quality aspects of tea after going through the Biswas' work. It is surprising to observe that after 1971, no serious attempt has been made (so far as our knowledge goes) to study the quality aspects in tea from statistical point of view. Not only the statistical approach, no other methodological approach has yet been investigated by the researchers to address this interesting problem of quality assessment. Only very recently, we have been informed that some studies have been initiated in the Sizukaya Tea Experimental Station, Japan, to address the problem of tea quality assessment using "Pattern Recognition" technique.**

However, several statistical studies have been made to understand the nature of sensory panel data specific to the food industry. The studies of Brockhoff et al. (1994) and Naes (1990) worth mentioning among others.

1.6 Data Description

In this section we discuss the data sets on which the whole study is based. There are nine sets of data, eight of which were provided by the Tea Research Association of India (TRA) and one data set was provided by the Tata Tea Ltd. (India).

Data Set 1 : Four sets of CTC samples were collected from the Tocklai garden in

four different years. All these samples were collected in the autumn flush period. The manufacturing/processing systems are same for all these samples and the biochemical and sensory analysis were conducted in the biochemistry and tea tasting laboratories of the Tocklai station. The samples are of sizes 25, 23, 25 and 21. The biochemical quality parameters are TF, TR, brightness (B), total liquor colour (TLC) and total soluble solids (TSS). The sensory analysis was done by the Taster of the experimental station and the samples evaluated in terms of quality and value on a 0-10 point scale. The basic statistics for these four data sets are presented in Table 1.1.

Data Set 2 : Under normal CTC processing system, 93 tea samples were studied in the Tocklai Station. An experienced Taster evaluated the tea samples in terms of 'strength' (S), 'quality' (Q), and 'valuation' (V). The scores were given on 0-10 point scale. We note here that by 'valuation ' we mean the overall quality here. The tea samples were plucked from the experimental garden in a particular flush period. It may be noted that in tea plantation there are four flush periods, namely, first flush, second flush, rain flush, and autumn flush. The quality of tea may vary over the flush periods. The biochemical parameters measured are TF, TR, B, TLC, and Total Soluble Solids (TSS). Some basic statistics for the chemical parameters and quality attributes are given in Table 2.1 and 2.2.

Data Set 3 : This data set was provided by the Tata Tea Ltd. (India). Fifty black CTC tea samples were collected from the Achbam Tea Estate of Assam (under Tata Tea Ltd.). The aliquots of each of these drier-mouth were tasted by an experienced Taster. The Taster evaluated the samples in terms of overall quality (V) and the scores were given on 0-5 point scale. From the data we observe that there are only three distinct scores, viz. 2.6, 3.00 and 3.40, assigned to the samples.

The biochemical analysis was done in the R & D center of Tata Tea Ltd., at Teok, Assam. The chemical parameters measured are TF, TR, TLC, and brightness (B). Some basic statistics for this data are presented in Table3.1.

Data Set 4 : This set of data contains 23 CTC samples, for which the biochemical

quality parameters measured are TF, TR, caffeine (C), crude fiber (CF) and Ash content (A). The sensory evaluation is made by a single Taster in terms of overall quality. The study was conducted in the experimental garden of Tocklai Experimental Station in the year 1997 as a part of the regular quality assessment study. The basic statistics for this data set are presented in Table 4.1 and Table 4.2. The Taster's scores (TS) are given on 0-10 point scale.

Data Set 5 : This sensory panel data is a part of the CTC manufacturing method development experiment, conducted in the year 1998 in the Tocklai Experimental Station. A panel of three Tasters evaluated the set of 14 samples in terms of S, Q and V. All the tasters scored the samples on a 0-10 point scale. Each Tasters made ten repeats for each sample, though only the mean of these repeats are provided to us. For certain reasons, the sensory panel data can not be presented here and only the basic statistics on these scores are given below. We can not specify the brand name of the samples also. The basic statistics are presented in Table 5.1.

Data Set 6 : For a set of 16 Tocklai released CTC clones a panel of five Tasters evaluated the samples in terms of S and Q in the year 1998. For each samples 10 repeats were made by each Taster. However, the mean of these repeats for each sample by each Taster was provided to us for study. The samples represent different clones though the identities of the products are not known to us. The basic statistics on the sensory scores are presented in Table 6.1.

Data Set 7 : This data is based on a trial experimental conducted in Tocklai to study the effect of a detergent, called Sumatotal, on the quality of CTC tea produced in upper Assam condition. This liquid detergent is used for washing the fermenting floor and green leaf processing machinery. The preconditioned leaf was rolled to pass through the CTC machine and then the leafs were spread with a thickness of 1 inch on the cement floor clean by Sumatotal liquid detergent to complete the fermentation. The processed tea samples were sent to a panel of four Tasters for sensory evaluation. The Tasters evaluated the 18 tea samples in terms of strength and quality on 0-10 point scale. The Tasters were of the

opinion that there was no significant development in the cup quality due to the use of the liquid detergent. We are studying only the experimental samples as the control samples are not available.

After processing the leafs using the detergent liquid, the samples were sent for the biochemical analysis. The biochemical parameters measured were TF, TR, TLC, WSS, and C. Some basic statistics about the data are given in the Table 7.1 and Table 7.2. S and Q denote scores on strength and quality respectively for the four Tasters.

Data Set 8 : A new process of manufacture of CTC tea has been developed at Tocklai station by modifying the sequences of manufacturing steps (Pal, Paul and Das, 1999). In this direction experiments were carried out in different commercial factories in various agro-climetic regions and the results confirmed higher percentage of finer grades and leaf appearances as well as improvement in cup quality. As per the modified method, when the plucked shoots were withered, rolled, fermented and then taken to CTC machine and dried instead of fermenting after the CTC cut, the product showed marked improvement in the cup quality and other quality attributes. The CTC manufacturing process has thereafter been adopted by different commercial gardens of Doors and Assam regions and outstanding performances have been observed in terms of auction price realization.

This data originates from the experimental results on the quality improvement of N-E India CTC tea by modifying the CTC processing system. The biochemical analysis and the sensory evaluations were done at the Biochemistry Laboratory and the Tea Tasting Department of Tocklai. The tea leaves were collected from the experimental garden of Tocklai as well as from different commercial gardens of Dooars and Assam regions. The experiments were aimed at modifying the CTC process to meet the market demand for higher percentages of Broken grades and leaf appearances.

Shoots plucked from the experimental garden were divided into two equal parts. One part of withered leafs was preconditioned in Rotorvane and passed through CTC machine three times and allowed to ferment on the floor at a thickness of spread of 1.25 cm. The fermentation time was kept between 1-10 min. to 1-30 min. depending on the temperature

and humidity. After the completion of fermentation the leaf was sent for drying. This is the conventional way of CTC processing (control samples). The other part of the leafs was passed through the same Rotorvane and was allowed to ferment on the floor with a thickness of spread of 3.75 cm. For a period varying between 1-10 to 1-30 hours. The preconditioned and fermented leaf was then passed through CTC machine three times and sent to dryer. Detail discussion on manufacturing method and the chemical properties is available in Pal et al. (2000). This is the experimental tea. The made tea samples were then drawn for different analysis. Nine control and experimental samples were considered for both biochemical and sensory evaluations.

Sensory analysis was done by a panel of four experienced Tasters in terms of overall value (V) on 0-10 point scale. The basic statistics on the chemical and sensory data are presented in Table 8.1.

Data Set 9 : This data represents the measurements on five biochemical quality parameters and the sensory scores are given by a panel of three Tasters on valuation (V) on a 0-10 point scale. The samples represent 30 Tocklai released CTC clones. There are two sets of information collected in the years 1997 and 1998. The chemical quality parameters studied are TF, TR, C, TLC and crude fiber (CF). The details on manufacturing methods and the system of chemical measurements are available in the Annual Scientific Report 1997 of TRA.

We note here that these CTC clones have different manufacturing/processing systems and naturally their chemical and quality characteristics vary. The clones are generally termed as CV1, CV2, etc. It is believed that the CV1 and CV2 are the best CTC clones among the 30 clones, in terms of quality and price realization. We have only one sample of each clone studied for each year. That is why we do not present the basic statistics on these two sets of CTC clone data. There is no meaning of giving information on means etc. taking over all the clones as each clone represents a particular type of tea under the CTC category.

1.7 Methodology

In this section we briefly introduce the methodologies used to study the quality aspects in tea. We explicitly state here that there may be various ways to approach the problem, we in our study limit the scope to only a few statistical investigations.

Primarily our aim is to study how closely the biochemical information can be functionally related to Taster's choice. For this we have associated the biochemical parameters with the choice made by a single Taster. The multiple regression techniques (partial regression analysis) have been applied to study the association on the basis of initial data diagnostics related to linearity etc. The performance of ridge regression technique and the robust technique like Minimum Absolute Deviation technique is compared with the ordinary least square (OLS) regression. Also the scores on different quality attributes given by a single Taster are associated with the biochemical information introducing Multiple Response Regression model.

The possible non-linearity in the data is studied separately using Box-Cox transformation model. The small sample estimation problem with the Box-Cox transformed linear models have been studied on the basis of a tea quality assessment data. The performance of non-linear least square estimation technique is compared with the maximum likelihood estimation (MLE) methods. A theoretical formulation of the Box-Cox type transformation model with measurement error in the response variable is also presented.

As a distinct approach to study the subjectivity of Taster's choices, we have obtained repeated observations on the quality assessment by taking more than one Taster on the same set of sample. Subjectivity in the sensory evaluation is studied introducing one-way and two-way variance components (VC) models with heteroscedastic formulation of variance-covariance matrix. The next step is to associate the sensory evaluations with the biochemical information after eliminating the bias due to Tasters. This is tried by developing regression models with repeated observations on the response variable. The random effects and mixed effects linear regression models are studied separately.

A review of the up-to-date literature on the studies made on different aspects of VC

models is presented in the second chapter. Special emphasis is given on the repeated measurement models. A detailed account of linear and non-linear models is given. Different aspects of longitudinal data analysis is discussed including approaches to data diagnostics. The analysis of variance (ANOVA), multivariate analysis of variance (MANOVA), generalized MANOVA (GMANOVA) model and estimation of variance components are discussed. This includes a section on profile analysis also. The general linear model for longitudinal data analysis is discussed in detail. The maximum likelihood (ML) and restricted maximum likelihood (REML) method of estimation are discussed along with the available algorithms for estimation. A detailed discussion on the testing aspects in general linear model is given. The review work is done keeping in line with the different repeated measurements models developed by us to study the quality aspects of tea.

In the third chapter, the association of different chemical quality parameters (specific to CTC tea) with a single Taster's scores on particular quality attribute is studied using regression techniques and the statistical significance of different chemical quality parameters in explaining the quality attribute(s) is observed. The multiple response regression model is applied to associate the biochemical quality parameters with a single Taster's choices on different quality attributes. The statistical aspects of measurement error with the response variable, when a single response is available, are discussed.

The possible non-linearity in the tea quality assessment data is studied using Box-Cox transformation model in the fourth chapter. Different computational problems associated with the Box-Cox models with small samples are discussed in detail.

The fifth chapter explicitly deals with the statistical analysis of sensory panel data. The statistical techniques useful to study the different possible variations in the sensory data are discussed in detail. The subjectivity of Tasters' choices is studied using one-way and two-way repeated measurement variance component models. Measurements made independently by different Tasters on quality attribute constitute the repeated observations on the sensory scores. Using random and mixed effects models the error variances associated with different Tasters' choices are estimated. The ANOVA and ML estimators of the heteroscedastic VC models are obtained. The statistical properties of the estimators

are studied.

The one-way error component linear regression model with repeated observations on the response variable is the theme of sixth chapter. The aim is obviously to associate the biochemical parameters with Tasters' scores after eliminating the bias associated with the sensory choices. The statistical significance of the chemical parameters in explaining the quality attribute(s) along with the error variances associated with the Tasters' scores may be studied using heteroscedastic formulation of the variance-covariance matrix. The ML and REML estimates of regression coefficients and the variance components are obtained employing a simple iterative algorithm. The statistical properties of the regression coefficients are also discussed.

A generalization of the linear random effect model is done introducing dummy variable to study the variations due to groups. This formulation is done to support a 'quality improvement experiment' for CTC tea, conducted in the Tocklai Experimental Station of TRA. A controlled experiment was carried out to develop a particular brand of CTC tea with better quality. A detailed discussion on the experiment is presented with original data. The experimental samples showed better quality as compared to the control sample.

The one-way error component regression model is appropriate specification when there is sufficient ground to believe that the tea samples under study are truly random samples representing a particular grade of tea and have the same intrinsic quality characteristic. However, in many situations, the quality characteristic may vary over samples. In such situations, apart from the variation due to Tasters, the sample specific effects should also be taken into consideration as an assignable source of variation. Keeping this aspect in view, we have discussed the two-way error component regression model in the seventh chapter.

The last chapter includes some discussions on several possibilities of technical studies to understand the association of biochemical quality parameters in tea with the Tea Tasters' sensory choices.

Table 1.1 : Summary statistics for the Data Set 1.

Parameters	Samples	Mean	S D	Minimum	Maximum
Quality	Sample = 25	5.48	1.89	2.00	8.00
	Sample = 23	5.83	1.70	4.00	8.00
	Sample = 25	5.72	1.51	4.00	8.00
	Sample = 21	4.78	1.40	2.00	8.00
Value	Sample = 25	7.84	0.85	6.00	9.00
	Sample = 23	8.00	0.80	7.00	9.00
	Sample = 25	7.88	0.78	7.00	9.00
	Sample = 21	7.44	0.71	6.00	9.00
TF	Sample = 25	0.80	0.06	0.67	0.93
	Sample = 23	0.80	0.06	0.72	0.94
	Sample = 25	0.82	0.05	0.73	0.91
	Sample = 21	0.83	0.09	0.66	0.94
TR	Sample = 25	9.88	1.49	7.06	12.02
	Sample = 23	9.97	0.94	7.87	11.39
	Sample = 25	9.21	0.59	8.06	10.61
	Sample = 21	10.04	0.72	8.94	11.12
B	Sample = 25	17.88	2.62	15.47	23.36
	Sample = 23	17.51	1.21	15.95	19.80
	Sample = 25	17.13	1.36	15.22	20.12
	Sample = 21	19.38	1.01	17.69	20.77
TLC	Sample = 25	3.08	0.29	2.75	3.90
	Sample = 23	2.87	0.22	2.34	3.30
	Sample = 25	2.95	0.18	2.55	3.29
	Sample = 21	3.04	0.24	2.34	3.30
TSS	Sample = 25	40.03	0.35	38.90	40.70
	Sample = 23	39.63	0.46	38.96	40.30
	Sample = 25	39.55	2.02	29.60	40.20
	Sample = 21	40.19	0.45	39.49	41.00

Table 2.1 : Summary statistics for Data Set 2.

Parameters	Mean	S D	Ranga	Correlation Metrix
V	2.90	0.23	2.60 - 3.40	1
TF	0.84	0.13	0.52 - 1.09	0.49 1
TR	12.32	1.78	8.07 - 16.13	0.53 0.57 1
TLC	4.33	0.62	3.03 - 5.85	0.56 0.83 0.54 1
B	15.97	1.51	13.61 - 19.67	-0.09 0.35 -0.23 -0.17 1

Table 3.1 : Summary statistics for Data Set 3.

Parameters	Mean	S D	Minimum	Maximum
S	5.46	1.67	2.00	8.00
Q	4.76	1.72	2.00	8.00
V	7.80	0.81	6.00	8.00
TF	0.81	0.06	0.66	0.94
TR	9.79	1.04	7.06	12.02
B	18.04	1.95	15.22	23.36
TLC	2.97	0.22	2.34	3.90
TSS	39.82	1.17	29.60	41.00

Table 3.2 : Correlation matrix for the Data Set 3.

S	1							
Q	0.73	1						
V	0.82	0.94	1					
TF	0.19	0.29	0.26	1				
TR	0.32	0.38	0.35	0.41	1			
B	-0.08	-0.07	-0.08	0.55	0.32	1		
TLC	-0.12	-0.04	-0.009	0.22	-0.08	0.32	1	
TSS	-0.10	-0.02	-0.02	0.19	0.13	0.29	0.15	1

Table 4.1 : Summary statistics for Data Set 4.

Variable	Mean	S.D.	Minimum	Maximum
TS	7.10	0.9472	5.50	9.20
TF	1.5339	0.2409	1.05	2.06
TR	12.45	1.0789	10.70	14062
CAF	3.6591	0.2769	2.99	4.06
CF	9.7948	0.7503	8.60	11.00
A	6.4913	0.2802	6.00	6.99

Table 4.2 : Correlation Matrix for Data Set 4

V	1					
TF	0.3765	1				
TR	0.2205	0.3742	1			
CAF	-0.1104	0.6978	-0.5929	1		
CF	-0.6233	-0.4198	-0.1541	0.1553	1	
A	-0.3316	-0.1032	-0.2971	0.3312	0.2342	1

Table 5.1 : Summary statistics for Data Set 5.

	Attributes	Taster 1	Taster 2	Taster 3
Mean	Strength	5.09	4.36	5.66
	Quality	6.50	6.34	6.33
	Value	7.96	7.55	7.00
S D	Strength	0.26	0.36	0.79
	Quality	0.78	0.63	0.63
	Value	0.31	0.77	0.96
Minimum	Strength	4.75	3.85	4.00
	Quality	5.28	5.43	5.00
	Value	7.57	6.00	6.00
Maximum	Strength	5.75	5.05	6.25
	Quality	7.75	7.25	7.25
	Value	8.42	9.00	9.00

Table 6.1 : Summary statistics for Data Set 6.

	Attribute	Taster 1	Taster 2	Taster 3	Taster 4	Taster 5
Mean	Strength	7.37	5.43	7.09	6.35	7.38
	Quality	7.36	5.38	7.31	4.22	7.41
S D	Strength	0.33	0.74	0.27	0.88	0.43
	Quality	0.35	0.68	0.43	0.34	0.44
Minimum	Strength	6.75	4.50	6.50	3.50	6.50
	Quality	6.75	4.50	6.50	6.50	6.50
Maximum	Strength	8.00	7.25	7.50	7.28	8.14
	Quality	8.25	7.00	8.00	8.00	8.00

Table 7.1 : Summary statistics for the Data Set 7.

	Mean	S D	Minimum	Maximum
S 1	6.00	1.37	4	8
S 2	7.33	1.19	6	10
S 3	6.33	1.97	4	10
S 4	8.44	1.10	6	10
Q 1	7.00	0.69	6	8
Q 2	7.78	0.65	7	9
Q 3	7.19	0.96	6	9
Q 4	8.28	0.57	7	9
TF	1.51	0.23	1.14	2.02
TR	12.78	1.01	10.90	14.08
WSS	45.02	0.86	43.20	46.58
TLC	5.53	0.28	5.05	5.98
C	4.28	0.42	3.67	5.37

Table 7.2 : Correlation Matrix for chemical parameters for Data Set 7.

	TF	TR	WSS	TC	C
TF	1				
TR	-0.48	1			
WSS	0.41	-0.19	1		
TLC	0.23	-0.35	0.41	1	
C	0.21	-0.04	0.31	0.37	1

Table 8.1 : Summary statistics for Data Set 8.

	Mean		S D		Minimum		Maximum	
	Exp.	Cont.	Exp.	Cont.	Exp.	Cont.	Exp.	Cont.
V 1	7.67	6.44	0.94	1.07	6	4	9	8
V 2	8.44	6.89	0.50	0.74	8	6	9	8
V 3	7.89	6.78	0.74	1.31	7	4	9	8
V 4	8.89	6.67	0.74	0.95	8	5	10	8
TF	1.66	1.45	0.24	0.54	1.17	1.14	2.02	1.60
TR	11.83	13.24	0.76	0.13	10.90	13.04	13.72	14.08
TLC	5.76	5.30	0.14	0.30	5.45	5.05	5.98	5.51
WSS	45.51	44.53	0.72	0.16	44.65	43.21	46.58	45.55
C	4.49	4.26	0.63	0.78	3.75	3.76	5.37	4.66
MO	3.43	3.46	0.32	0.54	2.97	2.45	.98	4.10

CHAPTER - 2

A REVIEW OF LITERATURE ON VARIANCE COMPONENT MODELS FOR REPEATED MEASUREMENTS

2.1 A Brief History of Developments in Variance Components Models

Scheffe (1956) and Anderson (1978, pp 11-25) gave a detailed account of the early history (1861-1949) of development of models and methods of estimating variance components. From their survey it appears that Legendre (1806) and Gauss (1809) implicitly dealt with fixed and random effects aspects of linear models in the field of astronomy.

The first known formulation of random effects model seems to that of Airy (1861), who simplicity used a variance Component (VC) model for one-way layout. Airy's study was on telescopic observations with respected measurements. It is note worthy that this earliest known use of VC model included unbalanced data structure. Airy assumed the following structure for the j^{th} observation on the i^{th} night.

$$y_{ij} = \mu + \alpha_i + e_{ij} \tag{2.1}$$

$$i = 1, 2, \ldots, a$$

$$j = 1, 2, \ldots, n_i$$

where μ is the true value, and the $\{\alpha_i\}$ and $\{e_{ii}\}$ are random effects. 'α_i' was termed as 'constant error'. The $\{e_i\}$ for fixed 'i' was explained as the errors about conditional mean $\mu + \alpha_i$. For the i^{th} night, he proposed the following estimate of error variance

$$\hat{\sigma}^2_{ei} = \sum_j (y_{ij} - \hat{y}_{i0})^2/(n_i - 1)$$

and the average of the square roots of the values was obtained as

$$\hat{\sigma}^2 = \left[\sum_i (\hat{\sigma}^2_{ei})^{1/2}/a\right]^2 .$$

The second user of a random effect model appears to be Chauvenet (1863). Coming into nineties, the major fundamental ideas on VCs are due to R.A. Fisher. He started with the basic paper is the theory of quantitative genetics (1918) where Fisher made inceptive use of the terms 'variance' and "analysis of variance". Following the genetics paper, Fisher's book (1925, Sec. 40) made a major contribution to the VC models. The basic approach of equating sum of squares (SS) from analysis of variance (ANOVA) to their expected values, and thereby obtaining a set of equations that are linear is VC's to be estimated, is due to Fisher.

Fisher did not use liner models to explain the ANOVA of designed experiments. In contrast L.C. Tippet (1931) not only classified the ANOVA method of estimating VC's from balanced data but also extended it to the two-way crossed classification, without interaction random model. Although Tippet considered an optimal design, the initial work on optimal sampling design through VC models (including higher order models) is due to Yates and Zacopany (1935). We note that at the early period of research on optimal sampling design, Cochran (1939) made substantial contribution.

Since linear models have now-a-days become an integral part of describing VC's, it is interesting to note that this had become widely accepted by 1939 : e.g. Neyman et al. (1935), Welch (1936), Deniels (1939). The specifications of the models were very much up-to-date in some cases : Welch utilized properties of χ^2-variates, Jackson (1939) assumed normality for random effects and error terms. The work of Deniels was significant as sampling variance of the VC estimates were derived for balanced date, up to the complexity of a 3 way crossed classification random effects model.

Cochran is the first statistician to discuss VC models for unbalanced data. Actually Cochran was not specifically concerned with the estimation of VC's, Winsor and Clarke (1940) actually did it for unbalanced data. The extension of general ANOVA method camee after 1940. Ganguly (1941) applied it to k-way nested classification random model

with interaction. In fact, Ganguly and Crumpt draw attention to a deficiency of the ANOVA method, namely that, depending on data it may produce negative estimates of VC. The aspects of negative estimates of variance components are discussed in details by Searle (1972). Crumpt also derived the sampling variance of the class of estimators for 1-way and 2-way crossed models. Wald (1941) considered confidence intervals for ratios of VCs.

The year from 1950 to 1969 brought major developments in the methods of estimating VCs, starting with important extensions of the methodology and ending with establishment of new methods based on maximum likelihood (ML) and minimum norm criteria. The first to mention is the Anderson-Bancraft's (1952) book which contains detail discussion on VCs. The book deals thoroughly with estimation of VCs from both balanced and unbalanced data for mixed and random models. This book is a milestone is the history of VC estimation. Details on the ML estimation of VCs would be discussed in separate section.

A landmark paper dealing with the difficult problem of how to use unbalanced data in VC models in due to Handerson (1953). This paper classified three different ways of using unbalanced data, from random and mixed models. All these three are applications of ANOVA method of equation SS to expected values. These three methods have came to be known as Handerson's Method I, II and III.

Keeping in view the question of optimality, several papers between 1956 to 1968 developed formulae for sampling variances of ANOVA estimates and of Handerson's method estimates in particular. The unbiasedness properties of different statistics for ANOVA estimates were first developed for balanced data by Graybill and Wortham (1956), and for unbalanced data by Scheffe (1959).

Whatever computability considerations on the part of different researches were there, the weaknesses of ANOVA estimators remained : negativity, lack of distributional properties and no useful way to compare different applications of ANOVA methodology. In light of that, the maximum likelihood (ML) estimation duly came to be considered as an better alternative. The initial effort in this line appears to lie with Crumpt (1947, 1951), who

dealt with 1-way classification. Herbech (1959) derived explicit ML estimators for certain balanced data models and also studied the non-negativity of VCs. The landmark paper for ML estimation in general is due to Hartley and Rao (1967), wherein a methodology was developed for a very wide class of models. But for few years there was an impediment to wide spread use of ML method due to computational complexities. Miller (1973, 1977) also worked on the ML estimation for balanced as well as unbalanced data. He showed that the ML equations for 2-way random model can be written in comparatively simple look, but these can not be solved analytically. He also studied the asymptotic properties of the estimators.

The study on restricted maximum likelihood estimation (REML) was initiated by W.A. Thompson (1962). He introduced the idea of maximizing that part of the likelihood which in invariant to the location parameters (the fixed effects) of the model. REML estimation for unbalanced data is due to Patterson and Thompson (1971). There is no denial of the great importance of Harville's (1977) effort to study the computational difficulties of ML as well as REML methods of estimation. This study reduced the computational confusions in variance components estimation.

The search for best linear unbiased estimated (BLUE) of VCs begun with Tounsend (1968), Harville (1969), Searle and Tounsend (1971). This is nothing but finding the minimum variance quadratic unbiased estimators of VCs, which is popularly known as Minimum Norm Estimation. The initial papers were quickly followed by La Motte's (1970, 71, 73a,b; 1976) work on minimum variance estimation and C.R.Rao's (1970, 1971a,b; 1972) study on minimum norm quadratic unbiased estimation (MINQUE). These estimators have, in some sense, a minimized generalized variance, derived from the minimization of Euclidean norm, which under normality, equates to a minimum variance property.

The MINQUE estimation demands no distributional assumptions. Also it does not involve any iteration procedure, just the solution of linear equations. The estimators are unbiased. We note that any MINQUE estimate is same as first round iterate from REML, using a priori value needed for MINQUE as the starting value for REML iteration. The importance of these connections between MINQUE and REML is discussed in the book of Rao and Kleffe (1988).

Lastly, we discuss in brief the developments made in the field of VC estimation through Bayesian approach. Some pioneering details is this line are due to Tiao and Tan (1965, 1966), Tiao and Box (1967), Hill (1965, 1967, 1970) and Culver (1971). Hill dealt with one-way classified data in balanced structure. Appropriate class of prior distribution on the components and/or their ratios is considered by Culver. There are only a few papers on Bayesian approach for unbalanced models. To cite a few works in our knowledge - Khuri and Sahai (1985), Gnot and Kleffe (1983), among others.

Coming on to the computational aspects, we note that much of the earlier estimation methods require approximation of integrals. The evaluation of very high dimensional integrals can be a computational problem (Smith, 1983). But there are numerous alternative methods available now-a-days for computing VC through Bayesian approach. Approximation of integrals, particularly arising from Bayesian hierarchical modeling, are treated in details by Tierney and Kadane (1986, 1989). Recent techniques include interesting work on application of Gibb's sampling (Gelfand and Smith 1990, Gelfand et. al., 1990) which can provide methods of obtaining estimates without doing the integration that the formal derivations dictate.

2.2 Linear Variance Component Models

Much of the early story of VC models revolves around the one-way classification which may be summarized as

$$y_{ij} = \mu + \alpha_i + e_{ij} \tag{2.2}$$

with $i = 1, 2, \ldots a$

$$
\begin{aligned}
Var(\alpha_i) &= \sigma^2, \quad Var(e_{ij}) = \sigma_e^2, \text{ all covariances zero} \\
j &= 1, 2, \ldots, n \quad \text{, for balanced data} \\
&= 1, 2, \ldots n_i, \text{ for unbalanced data.}
\end{aligned}
$$

(2.2) may be written in matrix notation as

$$\underset{\sim}{y} = (1_a \otimes 1_n)\, \underset{\sim}{\mu} + (1_a \otimes 1_n)\, \underset{\sim}{\alpha} + \underset{\sim}{e} \tag{2.3}$$

where a denotes the number of classes and n denotes the number of observations. Searle and Handerson (1979) and Anderson (1984) use extension of this random model. The dispersion matrices of $\underset{\sim}{y}, \underset{\sim}{\alpha}$ and $\underset{\sim}{e}$ are -

$$Var(\underset{\sim}{e}) = \sigma_e^2\, I_{an}, \quad Var(\underset{\sim}{\alpha}) = \sigma^2\, I_a, \quad and$$

$$Var(\underset{\sim}{y}) = \sigma^2(I_a \otimes J_n) + \sigma_e^2(I_a \otimes I_n) = I_n \otimes (\sigma^2 J_n + \sigma_e^2 I_n)$$

In some situation, adopting $Cov\ (\alpha_i, \alpha_i') = \rho\, \sigma^2)$ for $i \neq i'$ is reasonable. Then the general form is

$$Var(\underset{\sim}{\alpha}) = \sigma_\alpha^2[(1-\rho)\, I_a + \rho J_a].$$

The traditional fixed effects linear model may be written as

$$y = X\beta + e \tag{2.4}$$

where y is $(N \times 1)$ data vector, β is $(k \times 1)$ vector of fixed parameters occurring in the data, X is known $(n \times k)$ coefficient matrix and e is an error vector defined as $e = y - E(y) = y - X\beta$ and thus has $E(e) = 0$. Usually the dispersion matrix is $Var(e) = \sigma^2 I_N$. X is often matrix of zero and ones in 'no-regression' situation in which case it is known as 'incidence matrix'. But X may also include columns of regressors. To take care of all possibilities, X is called model matrix.

In VC models the random effects may be presented as $Z\ U$, where U is the vector of random effects that occur in the data and Z is the corresponding matrix, usually an

incidence matrix. Moreover, U can be partitioned into sub vectors. Incorporating the random component U into (2.3.4). We get a general from of model equation for a mixed model as

$$y = X\beta + Z\,U + e \tag{2.5}$$

where β and U represent for fixed and random effects respectively. We have $E(y) = X\beta$ and $E(y \mid U) = X\beta + Z\,U$. As mentioned above, partitioning Z and U, the mixed model may be represented as

$$y = X\beta + \sum_{i=1}^{q} Z_i\,U_i + e \tag{2.6}$$

Then we have, $V = Var(y) + Z\,D\,Z' + \sigma^2\,I_N = \sum_{i=1}^{q} \sigma_i^2\,Z_i\,Z_i' + \sigma^2\,I_N$. A useful extension is to consider $U_0 \equiv e$, $Z_0 = I_N$ and $\sigma_0^2 = \sigma_e^2$, and so have

$$y = X\beta + \sum_{i=0}^{q} Z_i\,U_i \quad \text{and} \quad V = \sum_{i=0}^{q} Z_i\,Z_i'\,\sigma_i^2.$$

The above formulations are due to Hartly and Rao (1967), who used these to great advantage for unbalanced data. Usually the following assumptions are made on the above formulation

$$Var(e) = \sigma^2\,I_N, \quad Cov(U_i, U_j') = 0 \;\forall\; i \neq j \quad \text{and} \quad Cov(U,\, e') = 0.$$

The zero covariances provides no opportunity to deal with situations where components of covariance would be appropriate. But many researchers have studied the components of covariances. The possible formulation in this line is outlined below :

i) Suppose u_{it} be an element of U_i for $t = 1, 2, \ldots q_i$. Suppose covariances between all pairs of elements of U_i are to be non-zero but covariances between different U's are to be zero; then

$$Cov\,(u_{it},\, u_{it'}) = d_{i,\,tt'} \quad and \quad Cov\,(u_{it}, u_{i't}) = 0\;\forall\; i \neq i'.$$

Hence $Var\,(U_i) = D_{ii}$ and $Var(U) \sum_i D_{ii}$. For example, in the intra-class correlation pattern discussed earlier, we have $D_{ii} = \sigma_i^2\,[(1-\rho_i)\,I_{qi} + \rho_i\,J_{qi}]$, which has $d_{i,tt} = \sigma_i^2$ and $d_{i,tt'} = \rho_i\,\sigma_i^2$.

ii) Another possibility is the covariance between effects of different random factors. In this situation one may assume,

$$Cov\ (u_{it}, u_{i't'}) = d_{ii',tt'} \quad \text{so that} Cov\ (\underset{\sim}{U}_i, \underset{\sim}{U}'_{i'}) = D_{ii'} = \sum_t \sum_{t'} d_{ii',tt'}$$

$$and \quad Var(U) = D = \sum_{ii'} D'_{ii}.$$

One possibility for D_{ii} is $D_{ii} = \sigma_i^2\ [(1 - \rho_{ii})\ I_{qi} + \rho_{ii}\ J_{qi}]$ and $D_{ii'} = \rho_{ii'}\ \sigma_i\ \sigma_{i'}\ J_{q_i \times q_i}\ \forall\ i \neq i'$, so that $Var(u_{it}) = \sigma_i^2\ \forall\ t = 1, 2, \ldots q_i$

$$Cov(u_{it}, u_{it'}) = \rho_{ii}\ \sigma_i^2\ \forall\ t \neq t' \quad \text{and} \quad Cov(u_{it}, u_{i't'}) = \rho_{ii'}\ \sigma_i\ \sigma'_i\ \forall\ t, t' \quad \text{and} \quad i \neq i'.$$

For error terms it is generally assumed that covariances between all pairs of error terms are zero. Also that all error terms have same variance σ_e^2. But one may assume some specific covariance structure for error component also. The situation of 'equi-correlated errors' may be considered. With diagonal elements σ_j^2 and off-diagonal elements $\rho\sigma_j\ \sigma_k$ in the variance-covariance matrix V, we may write $V = s[(1 - \rho)I_a + \rho\ J_a]s$, where $s = diag(\sigma_1, \sigma_2 \ldots \sigma_a)$. Such a structure may be considered when all measurements are made at about the same time. Such situation are encountered in 'split-plot' type set-up. Again when the measurements on an individual subject have been made in sequence over time, the errors may be correlated. A widely used time series model is the autoregressive process

$$e_j = \rho\ e_{j-1} + u_j\ , \text{ for } \quad j > 2$$

where ρ is the regression parameter and u_j, are innovation errors usually assumed to be $N(0, \sigma^2)$, each independent of the past. Again when the data are unequally spaced over time, some researchers propose the "Markov correlation structure" for error component.

A useful review in this field with many references is due to Mukherjee (1976). Rao (1967) considers least square (LS) estimators with unstructured and autoregressive forms of V. Webb (1973) applied the non-stationary autoregressive model. Beach and MacKinnon (1978) studied the ML estimation aspects in the autoregressive case. Wilson et

al. (1981) considered, among other things, MLE for the model $y_i = \mu 1_r + e_i$ with $V_i = \sigma^2 J_r + \Sigma$, Σ having the autoregressive form discussed above. Azzalini (1984) added the regressors (covariates) to this scheme. To name few other contributors in this field are Kenward (1985), Rochan et al. (1989), Jennrich and Schluchter (1986).

2.2.1 ANOVA Estimators of Variance Components

We now discuss the ANOVA estimation of four different VC models which have been used extensively by bio-statisticians. For the one-way random model discussed earlier, the ANOVA table and the estimators are given below :

Source	d.f.	S. S
A	$a-1$	$SSA = \sum_i n\,(\bar{y}_{i0} - \bar{y}_{00})^2$
Residual	$a-1$	$SSE = \sum_i \sum_j (y_{ij} - \bar{y}_{i0})^2$
Residual	$an-1$	$SST = \sum_i \sum_j (y_{ij} - \bar{y}_{i0})^2$

$\sigma_e^2 = MSE,$

$\sigma_e^2 = (MSA - MSE)/n$

The two-way nested random model is

$$\begin{aligned} y_{ijk} &= \mu + \alpha_i + \beta_{ij} + e_{ijk} \\ i &= 1,2,\ldots a;\ \ j = 1,2,\ldots b;\ \ k = 1,2,\ldots n. \end{aligned}$$

The ANOVA table is

Source	d.f.	S. S
A	$a-1$	$SSA = \sum_i bn\,(\bar{y}_{i00} - \bar{y}_{000})^2$
B Within A	$a(b-1)$	$SSB:A = \sum_i \sum_j n\,(\bar{y}_{ij0} - \bar{y}_{i00})^2$
Residual	$ab(n-1)$	$SSE = \sum_{ijk}(y_{ijk} - \bar{y}_{000})^2$
Total	$abn-1$	

$\hat{\sigma}_e^2 = MSE$

$\hat{\sigma}_\alpha^2 = (MSA - MSB:A)/bn$

$\hat{\sigma}_\beta^2 = (MSB:A - MSE)/n$

The two-way crossed, with interaction mixed model is

$$y_{ijk} = \mu + \alpha_i + \beta_j + \lambda_{ij} + e_{ijk}$$

with α_i's fixed. The ANOVA table is

Source	d.f.	S. S
A	$a-1$	$SSA = \sum_i bn\,(\bar{y}_{i00} - \bar{y}_{000})^2$
B	$a(b-1)$	$SSB = \sum_j an\,(\bar{y}_{0j0} - \bar{y}_{000})^2$
AB	$(a-1)(b-1)$	$SSAB = \sum_{ij} n\,(\bar{y}_{ij0} - \bar{y}_{i00} - \bar{y}_{0j0} + \bar{y}_{000})^2$
Residual	$ab(n-1)$	$SSE = \sum_{ijk} (y_{ijk} - \bar{y}_{ij0})^2$
Total	$abn-1$	

$\hat{\sigma}^2_\beta = (MSB - MSAB)/an$

$\hat{\sigma}^2_\lambda = (MSAB - MSE)/n$

$\hat{\sigma}^2_e = MSE.$

Again the 2-way crossed, no interaction mixed model equation is

$$y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk}$$

with α_i's fixed. The ANOVA table and the estimators of VC are

Source	d.f.	SS
A	$a-1$	$SSA = \sum_i bn(\bar{y}_{i00} - \bar{y}_{000})^2$
B	$b-1$	$SSB = \sum_j an(\bar{y}_{0j0} - \bar{y}_{000})^2$
Error	$abn-a-b$	$SSE = \sum_{ijk}(y_{ijk} - \bar{y}_{i00} - \bar{y}_{0j0} + \bar{y}_{000})^2$
Total	$abn-1$	

$\hat{\sigma}_\beta^2 = (MSB - MSE)/an$

$\hat{\sigma}_e^2 = MSE.$

2.2.2 Studies with Repeated Measurements

We now turn our discussion on the situation where repeated measurements are available on both the covariates (regressors) and the response variate. It means that both y and X have errors and repeated measurements are available on both of them. Such problems have been discussed extensively by Madansky (1959) and Cochran (1968). If we have N_i observations on each (X_i, y_i) with

$$y_{ij} = y_i + v_{ij} \qquad i = 1, 2, \ldots n$$
$$x_{ij} = X_i + u_{ij} \qquad j = 1, 2, \ldots N_i$$

and if the usual assumptions of independence are made, then we can perform ANOVA on the X's and y's and obtain the estimates of β.

Lord (1960), Dagracie and Fuller (1972), proposed an estimate of the following functionally related covariance model.

$$\begin{aligned} y_{ij} &= \alpha + \lambda_i + \beta x_{ij} \\ & \qquad i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, r \end{aligned}$$

where n is the number of treatments and r is the number of repetitions. λ_i is the i^{th} treatment effect with $\sum \lambda_i = 0.x_{ij}$ and y_{ij} are observed with errors following a bivariate normal form with zero means. For known estimates of variance components, they developed the estimators of β that are unbiased to $O(r^{-1})$ where r is the number of observations on each treatment.

Ord (1969) assumed a model where replicated observations (only two) are possible for fixed true vales of the variables and obtained the ML estimators for the functional relationship

$$\left.\begin{aligned} y_i &= \alpha + \beta x_i \\ \lambda_i &= X + \partial_{ij} \\ \eta_i &= y_i + \epsilon_i \end{aligned}\right\} \begin{aligned} i &= 1, 2, \ldots n \\ j &= 1, 2, \end{aligned}$$

with usual assumptions. This may be relevant when observations are based on two independent situations. Some good discussions in the field of measurement errors with repeated observations on both the covariates and response are made in the book of Carrol et al. (1995). A detail statistical study on the error-in-variables is due to Pal (1981).

Many of these models can be generalized in the longitudinal set up as discussed below.

2.3 Some Discussions on Longitudinal Studies

Longitudinal studies represent one of the principal research strategies employed in biomedical and social science research. The defining characteristic of a longitudinal study is that individuals are measured repeatedly through time. This is obviously in contrast to cross-sectional studies, in which a single outcome is measured for each individual. In longitudinal data the response of each individual (sample) is observed on two or more occasions.

Longitudinal designs are uniquely suited to the study of individual change over time, including the effects of development, aging and other factors that effect change. Longitudinal studies typically have unbalanced designs, missing data, time-varying covariates, and other characteristics that make standard multivariate procedures (e.g. MANOVA etc.) inapplicable. The major advantage of longitudinal study is its capacity to separate what in the context of population studies are called cohort and age effects.

The defining feature of a longitudinal data is repeated observation on individuals allowing direct study of changes. Longitudinal data require special statistical methods because the set of observations on one subject tends to be interrelated generally. The research of eighties focused on the development of statistical methods that not only consider the inter-correlation of serial measurements but also accommodate the complexities of typical longitudinal data sets. Ware(1985) viewed the analysis of serial measurements as a univariate regression analysis of responses with correlated errors. He discussed more flexible approaches to modeling and parameter estimation. He argued that the repeated measures designs may be regarded as a subset of longitudinal designs. The methods developed for longitudinal designs can be directly applied to data collected in the repeated measurers setting.

The issue for accounting correlation also arises when analyzing a single long time series of measurements. Diggle (1990) discusses time series analysis in the biological sciences. Analysis of longitudinal data tends to be simpler when subjects can usually be assumed normal. However, in many situations the non-normal patterns are observed, which demand special statistical treatments. The inferences from longitudinal studies can be made more robust to model assumptions than those from time series data, particularly to assumptions about the nature of the correlation.

2.3.1 Approaches to Longitudinal or Repeated Measures Data Analysis

When we have single observation on each unit, then we are confined to modeling the population average of response y, called the marginal mean response. In this case we have no other choice. But with repeated measurements, there are several different approaches

that can be adopted. A simple strategy may be to :

(i) Reduce the repeated observations into one or two summaries;

(ii) Analyze each summary variable as a function of covariates X.

Such an approach in adopted in Pal and Paul (1997), where a summary measure of quality scores on different tea samples are obtained from the repeated observations on quality scores given by Tea Tasters. Than these summary scores are regressed on the corresponding measures of biochemical quality parameters and the coefficient estimates are obtained. We may call this approach a two stage or derived variable analysis. But it is worth noting that this approach is less useful if the most demanding explanatory variable change over time.

An alternative to the above approach may be to model the individual y_i, in terms of x_{ij}. The first approach is to model the marginal mean as in cross sectional studies (Diggle et al. 1995, p.18). Since repeated observations are likely to be dependent, the marginal analysis must include assumptions about the form of correlation. For example, in the linear model we may assume $E(y_i) = X_i\beta, Var(y_i) = V_i(\alpha)$, where β and α must be estimated. This approach carry the advantage of separately modeling the mean and covariance. Valid inference about β can sometimes be made even for incorrect form of $V(\alpha)$.

A second approach, the random effects model assumes that correlation arises among repeated responses because the regression coefficients vary across individuals. Here, we model conditional mean of y_{ij}, given β_i, by $E(y_{ij}/\beta_i) = x'_{ij}\ \beta_i$. It may be noted that fixed effects model is a appropriate specification if focus is on specific set of n individuals (samples). Inference is this case is conditional on the particular n individuals. Again random effects model is appropriate specification if we are drawing individuals randomly from large population.

Another approach, known as 'transition model' (Ware et al. 1988) focuses on the conditional expectation of y_{ij} given past outcomes, $y_{i,j-1}, \ldots y_{i1}$. Here we specify regression model for the conditional expectation, $E(y_{ij}/y_{i,j-1}, \ldots y_{i1}, x_{ij})$, as an explicit function of

x_{ij}, and of the past responses. An example of such model is the logistic regression model for binary data

$$log \frac{P_r(y_{ij} = 1/y_{i,j-1}, \ldots y_{i1}, x_{ij})}{1 - P_r(y_{ij} = 1/y_{i,j-1}, \ldots y_{i1}, x_{ij})} = x'_{ij}\beta + \alpha y_{ij-1}.$$

Transition model of this type combine the assumptions about the dependence of y on X and the correlation among repeated ys into a single equation.

In each of the three approaches discussed above, we model both the dependence of y on X and the autocorrelation among ys. With cross sectional data, only the dependence of y on X need to be specified; there is no correlation.

Thus longitudinal data analysis or repeated measurements problems may be partitioned into two groups :

(i) where regression of y an X is the point of interest and the number of observation (n) is greater than the number of repetitions (r).

(ii) problems where the correlation among repeated observation are of prime interest or when n is small.

2.3.2 Data Diagnostics

The longitudinal data analysis has two components that operate side by side : exploratory and confirmatory analysis. Exploratory Data Analysis (EDA) comprises techniques to visualize the data patterns. Confirmatory analysis is obviously technical in nature, weighing evidence in data for or against hypothesis.

Most longitudinal analysis address relationship of response with explanatory variables, often including time. So a scatter plot of responses against an explanatory variable may be the basic display. Smoothing techniques are discussed in literature that highlights the typical response as a function of explanatory variable without reliance on specific parametric models. Smoothing splines, kernel estimators, and the robust method 'lowess' are reviewed in Diggle et al. (1995).

We may also explore the correlation structure for degree of association in repeated measurements. To remove the effects of X, we first regress the response, y_{ij} on x_{ij}, to obtain residuals $e_{ij} = y_{ij} - x'_{ij}\,\beta$. With data collected at a fixed number of equally spaced points, the correlation can be studied using scatter plot matrix in which e_{ij} is plotted against e_{ik} for all $j < k = 1, \ldots r$. When each scatter plot in the matrix appears like a sample from the bivariate normal distribution, we may summarize the association with a correlation matrix, comprised of a correlation coefficient for each plot.

The best sources for studies in EDA are the books by Tukey (1977) and by Mosteller and Tukey (1977). Background information on graphical methods in statistics can be found in Chambers et al. (1983).

2.3.3 Repeated Measures ANOVA

The repeated measures ANOVA can be regarded as a initial attempt to a single analysis of a complete longitudinal data set. These aspects are discussed in detail earlier. Here we outline the 'split-plot' type approach, which was adopted by researchers in different agricultural studies. The underlying model may be presented as

$$y_{ijh} = \beta_h + \alpha_{hj} + \lambda_{hi} + e_{ijh} \quad i = 1, 2, \ldots, n, \;\; j = 1, 2, \ldots, r, \;\; h = 1, 2, \ldots g$$

where y_{ijh} denotes the j^{th} repeated observation for i^{th} sample within h^{th} treatment group. β_h represent the main effects for treatments and α_{hj} interaction between treatment and repetition with the constraint $\sum_j \alpha_{hj} = 0$. λ_{hi} are mutually independent random effects. e_{ijh} are mutually independent measurement errors. We have, $E(y_{ijh}) = \beta_h + \alpha_{hj}$. Under the assumptions $\lambda_{hi} \sim N(0, \sigma^2_\lambda)$, $e_{ijh} \sim N(0\sigma^2)$, the resulting distribution of $\underset{\sim h}{y} = (y_{i1h}, y_{i2h}, \ldots y_{irh})$ is multivariate normal with covariance matrix $V = \sigma^2\, I + \sigma^2_\lambda\, J$, where J is the matrix of ones. This implies a constant correlation $\rho = \sigma^2_\lambda/(\sigma^2_\lambda + \sigma^2)$, between any two observations on the same sample.

The split-plot ANOVA for the above model is presented in the following Table. In the Table $n = \sum_h n_h$ denotes the total number of units. We note here that split - plot

ANOVA requires a complete data or balanced data. But we may analyze incomplete data under split-plot model by general likelihood based approach (Diggle et al. 1995).

Table : Split-Plot ANOVA Table

Source	SS	d. f.
Between treatment	$BTSS_1 = n\sum_h n_i(\bar{y}_{h..} - \bar{y}_{...})^2$	$g-1$
Whole plot residual	$RSS_1 = TSS_1 - BTSS_1$	$n-g$
Whole plot total	$TSS_1 = n\sum_h \sum_i (\bar{y}_{hi.} - \bar{y}_{...})^2$	
Between repeats (time)	$BRSS_2 = r\sum_j (\bar{y}_{..j} - \bar{y}_{...})^2$	$(r-1)$
Treat x repeat interaction	$ISS_2 = \sum_j \sum_h r_h(\bar{y}_{hi.j} - \bar{y}_{...})^2 - BTSS_1 - BTSS_2$	$(g-1) \times (r-1)$
Split plot residual	$RSS_2 = TSS_2 - ISS_2 - BTSS_2 - TSS_1$	$(n-g) \times (r-1)$
Split plot total	$TSS_2 = \sum_h \sum_i \sum_j (\bar{y}_{hij} - \bar{y}_{...})^2$	$(nr-1)$

The split-plot ANOVA contains strong assumptions about the covariance structure, and hence a model based analysis under the assumed uniform correlation structure achieves a lot. We may adjust for missing values and also allow a structured linear model for the mean response profiles. But it is worth noting that although ANOVA methods are useful in particular circumstances, they do not constitute a general viable approach to longitudinal or repeated measurements analyses.

2.4 MANOVA and GMANOVA Models for Repeated Measures

2.4.1 MANOVA Models

The multivariate analysis of variance model (MANOVA) consists of p different response variables which are observed for each of n experimental units or subjects. The responses can be distinct variables or repeated measurements of one variable, or repeated measurements of a set of variables.

In such situations one may opt for the multiple design multivariate (MDM) liner model or the Zelner's seemingly unrelated regressions (SUR) model. The MDM linear model has applications to the multivariate analysis of repeated measures and crossover experiments. In many psychometric studies such approach is adopted.

Let $\underset{\sim i}{y} = (y_{i1}, y_{i2} \ldots y_{ip})'$ be p measurements on i^{th} individual $(i = 1, 2, \ldots, n)$, which is treated as a single vector multivariate observation. The general model on which the analysis would be based is $\underset{\sim i}{y} = \underset{\sim i}{\mu} + \underset{\sim i}{e}$ for individual i, corresponding to (2.2), where the errors $\underset{\sim i}{e}$ are independent with mean $\underset{\sim i}{0}$ and covariance matrix $V(\underset{\sim i}{e}) = \Sigma$. Thus Σ is $(p \times p)$ with $(j,k)^{th}$ element $Cov(e_{ij}, e_{ik}) = \Sigma_{ik}$. We note that $V(\underset{\sim i}{y}) = V(\underset{\sim i}{e}) = \Sigma$ is same for all i.

It is convenient to express the MANOVA model in matrix notation as

$$y = A\beta + e \tag{2.7}$$

where $y = [\underset{\sim 1}{y'}, \underset{\sim 2}{y'} \ldots \underset{\sim n}{y'}]'$, $A = [\underset{\sim 1}{a'}, \underset{\sim 2}{a'} \ldots \underset{\sim n}{a'}]'$ and $e = [\underset{\sim 1}{e'}, \underset{\sim 2}{e'} \ldots \underset{\sim n}{e'}]'$ are $(n \times p), (n \times q)$

and $(n \times p)$ matrices respectively. Generally A is assumed to be of full rank q so that the increase exists. It is generally assumed that $\underset{\sim}{e} \sim N_p(\underset{\sim}{0}, \Sigma \otimes I_n)$.

Clearly the MANOVA model (2.7) consists of p distinct, but correlated, univariate linear models each with the same between - subjects design matrix A. The ML estimator of β and dispersion matrix Σ are

$$\hat{\beta} = (A'\,A)^{-1}\,A'\,y \quad \text{and} \quad \hat{\Sigma} = Q/n,$$

where $Q = y'[I_n - A\,(A'A)^{-1}A']\,y$ is the error SS and cross product matrix (Anderson, 1984).

In multivariate regression study, we have a system of p separate regression models which are correlated. As is known in MANOVA framework, the same set of regressors is used for each of the p response variates.

The multivariate analysis of covariance (MANCOVA) is well described in Anderson's book (1984). The typical MANCOVA model is mixture of A and the parameter matrix β is partitioned into groups according to the need of the experiment. In fact, the theory for the comparison of different group effects are well developed in the literature of multivariate analysis and 'Profile Analysis' is one of the popular techniques in such situations.

2.4.2 Profile Analysis

Profile analysis pertains to situations where a battery of p-treatments are assigned to two or more groups of subjects. All responses must be expressed in similar units. For different groups the responses should be independent. The basic question is whether the mean vectors are same or not. In profile analysis, the question of equality of mean vectors is divided into specific possibilities.

Suppose $\underset{\sim}{\mu}_1'$ and $\underset{\sim}{\mu}_2'$ are mean responses of p-treatments for two populations. We can formulate the question of equality in a stagewise fashion.

(i) Are the profiles parallel ?

i.e., $H_0 : \mu_{1i} = \mu_{1\,i-1} = \mu_{2i} - \mu_{2\,i-1} \quad \forall\, i$

or equivalently $H_0 : C\underset{\sim 1}{\mu} = C\underset{\sim 2}{\mu}$ where C is $(p-1) \times p$ contrast matrix.

(ii) If the profiles are parallel do they coincide ?

i.e., $H_0 : \mu_{1i} = \mu_{2i} \quad \forall\, i$

or, $H_0 : 1'\underset{\sim 1}{\mu} = 1'\underset{\sim 2}{\mu}$

(iii) If the profiles are coincident, are all the means equal to same constant?

i.e., $H_0 : \mu_{11} = \mu_{12} = \ldots = \mu_{21} \ldots = \mu_{2p}$

or, $H_0 : C\underset{\sim}{\mu} = 0$

In all these cases the related T^2 statistics are described by Johnson and Wichern (1992).

In repeated measurements design, if we are interested in comparing the mean - effects (for two groups only), then we may represent the cell mean model as

$$\underset{\sim ij}{Y} = \underset{\sim j}{\mu} + \underset{\sim ij}{e} \;,\; j = 1, 2 \tag{2.8}$$

where $\underset{\sim j}{\mu} = [\mu_{j1}, \mu_{j2} \ldots \mu_{jp}]'$ is the mean vector response for j^{th} group. Alternatively the cell mean model may be written in terms of the one-way MANOVA model(2.7), where $y = \left[\underset{\sim 11}{Y} \ldots \underset{\sim n_1 1}{Y} : \underset{\sim 12}{Y} \ldots \underset{\sim n_2 2}{Y}\right]' = \left[\underset{\sim 1}{Y} \ldots \underset{\sim n}{Y}\right]'$ is the $(n \times p)$ observation matrix with $n = n_1 + n_2$ and $e = \left[\underset{\sim 1}{e} \ldots \underset{\sim n}{e}\right]'$ is the corresponding random error matrix. The design matrix A may be presented as

$$A = \begin{bmatrix} j_{n_1} & 0_{n_1} \\ 0_{n_2} & j_{n_2} \end{bmatrix}, \beta = \begin{bmatrix} \beta_{11} & \ldots & \beta_{1p} \\ \beta_{21} & \ldots & \beta_{2p} \end{bmatrix} = \begin{bmatrix} \underset{\sim 1}{\mu} \\ \underset{\sim 2}{\mu} \end{bmatrix}.$$

Here β represents the cell means. The primary hypotheses of interest may be written as

$$H_1 : C_1\, \beta\, U_1 \;=\; 0 \text{ (equal group effects)}$$

$$H_2 : C_2\,\beta\,U_2 \;=\; 0 \text{ (equal time effects)}$$

$$H_3 : C_3\,\beta\,U_3 \;=\; 0 \text{ (no groups} \times \text{time interaction)}$$

where $C_1 = [1\ -1]$, $U_1 = j_p$, $C_2 = [1\ 1] = C_3$, U_2 & U_3 are contrast matrices. For all these hypotheses to be tested, the UMP invariant tests are given by T^2 statistics. The testing aspect is well described by Vonesh and Chinchilli (1997).

2.4.3 The GMANOA Model

The generalized multivariate analysis of variance (GMANOVA) is another important aspect of repeated measurement design. It is a linear regression type approach, and is more flexible than ANOVA or MANOVA. Pathoff and Roy (1964), Roy (1967) introduced the GMANOVA model.

A GMAOVA model for balanced and complete data may be presented as

$$y = A\beta X + e \tag{2.9}$$

where y is $(n \times p)$ response matrix, β is $(q \times t)$ unknown parameter matrix, X is $(t \times p)$ within-subject design matrix with full rank $t(\leq p), e$ is $(n \times p)$ random error matrix. The distributional assumption for this model is same as MANOVA model when X is square, i.e. $t = p$, because X being invertible leads to $E(y\,X^{-1}) = A\beta$.

Here the elements of X may be considered as regressors, but not necessarily. X may be constructed so as to contain binary indicator variable in order to model the within subjects main effects and interaction (if any). The difference between MANOVA and GMANOVA models for repeated measures design is that the former requires all within subject effects to included in the model $(t = p)$, whereas the later does not $(t \leq p)$. Thus GMANOVA model with $t < p$, may lead to more efficient estimation as to contain fewer parameters.

The ML estimator for β and Σ under GMANOVA set up is given by

$$\begin{aligned}\widehat{\beta} &= (A'\,A)^{-1}\,A'\,y\,Q^{-1}\,X'(X\,Q^{-1}\,X')^{-1}\\ \widehat{\Sigma} &= \frac{1}{n}\left[Q + W'\,y'\,A(A'A)^{-1}\,A'\,yW\right],\end{aligned}$$

where $Q = y'\,[I_n - A\,(A'A)^{-1}\,A']\,y$,

$W = I_p - Q^{-1}\,W'(X\,Q^{-1}\,X')^{-1}\,X$.

Vonesh and Chinchilli (1997) give the proof of these results. Grizzle and Allen (1969) developed a goodness of fit test for the within units design matrix X in GMANOVA model. Khatri (1973) and Lee (1991) described different tests for certain variance structures within the GMANOVA model. Khatri developed different LR tests. Lee developed LR tests when Σ has autocorrelation structure. Puri and Sen (1985) described the use of rank statistic for GMANOVA model.

2.5 The General Linear Model for Longitudinal Data

In seventies the analysis of longitudinal data generally used a split-plot type model. These models required an assumption of equal variance covariance for repeated measurements, even though MANOVA approach (Cobe and Grizzle, 1966) and growth curve analysis methods already existed. In growth curve analysis, the expected response is modeled as a continuous function of time. Techniques for analyzing incomplete data from general (unstructured dispersion matrix) multivariate normal population were developed (Dempster et al., 1977). However these methods were found not suitable for clinical and some other biomedical studies, specially when the number of measurements on a subject is large relative to the number of subjects.

The work having major impact of clinical trials with repeated measures designs come from Laird and Ware (1982). Based on the work of Harville (1977), they developed ML and restricted maximum likelihood (REML) procedures for analyzing a general mixed effect model for repeated measurements.

Let $\underset{\sim}{\alpha}$ denote $(p \times 1)$ vector of population parameters and X_i be a known $(n_i \times p)$ design

matrix linking $\underset{\sim}{\alpha}$ to y_i. Let $\underset{\sim_i}{b}$ denotes a $(k \times 1)$ vector of unknown individual effects and Z_i a known $(n_i \times k)$ design matrix linking $\underset{\sim_i}{b}$ to $\underset{\sim_i}{y}$. For measured, multivariate normal data, Laird and Ware proposed the following model :

Stage 1 : For each individual i,

$$\underset{\sim_i}{Y} = X_i \underset{\sim}{\alpha} + Z_i \underset{\sim_i}{b} + \underset{\sim_i}{e} \tag{2.10}$$

where $\underset{\sim_i}{e} \sim N(0,\ R_i)$, R_i being a positive definite covariance matrix. At this stage, $\underset{\sim}{\alpha}$ and $\underset{\sim_i}{b}$ are assumed fixed, and $\underset{\sim_i}{e}$ are assumed to be independent.

Stage 2 : The $\underset{\sim_i}{b} \sim N(0, D)$ independently of each other and of the $\underset{\sim_i}{e}$. D is $(k \times k)$ positive definite covariance matrix. The population parameters $\underset{\sim}{\alpha}$ are treated as fixed effects.

Marginally, $\underset{\sim_i}{Y} \sim N(X_i\ \underset{\sim}{\alpha},\ R_i + Z_i\ D\ Z_i')$. Further simplification of the model is when $R_i = \sigma^2 I$. This allows us to write likelihood as the product of marginal densities of $y_1, y_2, \ldots y_n$. D and R_i are assumed to have same structures so that their elements can be written as a function of parameters on a lower dimensional space.

When one or more columns of X_i are function of time points and $X_i = Z_i$, the model serves as a growth curve model. Other columns present either the overall mean or changing covariates (regressors).

The model can accommodate any missing date pattern. Jennrich and Schluchter (1986) illustrated this model with different structures of $\Sigma_i = R_i + Z_i\ D\ Z_i'$ and proposed computing algorithms for ML and REML estimates and the corresponding LR tests under normality. They considered independence, compound symmetry, random effects, AR(1) and unstructured models for Σ_i.

We note here that no systematic efforts have been made to suggest a practical structure for the covariance matrix Σ_i for a repeated measurement design. Test of goodness of fit of a model with particular covariance structure is difficult, but its asymptotic LR test is, in general, sensitive to the departure from multivariate normality.

The Laird-Ware model has the advantage of combining both one stage and tow-stage regression models (Crowder and Hand, 1995). That is why, it is a more flexible model

as compared to the general models proposed by earlier researchers. As the Laird-Ware model also include two stage regression format, this may easily be reduced to a random coefficient growth curve model. Random coefficient regression model include growth curve models as discussed by C.R. Rao (1965), Swamy (1970), Lindsey and smith (1972), Fearn (1975) and many others.

Much of the classical work on two stage models concerns the fitting of polynomial curves to animal growth measurements over time. Here the design matrix X has j^{th} raw $(1, x_j, x_j^2, \ldots x_j^{q-1})$ or the orthogonal polynomials version of this. Also in the literature (see Vonesh and Chinchilli, 1997) X_i has generally been equal to Z_i or $Z_i\ A_i$ for some A_i. This arises from $\underset{\sim}{Y}_i = Z_i\ \underset{\sim}{\alpha}_i + \underset{\sim}{e}_i$ at the first stage, and than $\underset{\sim}{\alpha}_i$ having distribution $N(A_i\ \alpha,\ D)$ at the second stage. Thus $E(y_i) = X_i\ \alpha$, with $X_i = Z_i\ A_i$ and $Var\ (y_i) = Z_i\ D\ Z_i' + R_i$.

Heitjan (1991) has proposed same generalization of growth curves for repeated measures design basing on the logistic growth curves proposed by Helder (1961). A good discussion on linear and non linear growth curves is available in Lindsey's book (1993).

2.5.1 Studies With Non-Normal Distributions

Recently mixed effects models for a distribution from the exponential family have received considerable attention. Beitler and Landis (1985) considered a mixed effects model with no covariates, directly for a binary response. They computed the VC from the quadratic forms from the conventional ANOVA table as one would obtain from normal data. However, the validity and efficiency of these estimates are questionable.

For analyzing a longitudinal data satisfying a distribution from the exponential family, several methods have been developed. The empirical generalized least squares (EGLS) procedure developed by Koch et al. (1977) exploits full multinomial structure in computing the dispersion matrix for the estimates. Although computationally more complicated, the generalized estimating equations (GEE) approach for marginal models, proposed by Liang and Zeger (1986), has certain advantages. The population averaged parameters are modelled as functions of covariates in marginal models. The main advantage of GEE is that it accomodates continuous time dependent or time independent covariates. Also,

the dispersion matrix can be modeled in terms of fewer parameters than the number of parameters in an unstructured dispersion matrix and thus the consequence of sparse data can be avioded. The GEE approach does not attempt to model the joint distribution of the repeated measurements. The marginal distribution at each time point is modeled as a function of covariates. Allowing a working correlation matrix among the subject responses the regression parameters and their dispersion matrix are estimated. These estimators are consistent as long as the population means are correctly specified. Otherwise, there is some loss of efficiency. This procedure is a multivariate extension of quasi-likelihood and is not a likelihood based procedure. Prentice and Zhao (1991) has contributed to extend this theory.

Another approach for analyzing random effects model is to use the conditional likelihood given sufficient statistics for the subject effects. Diggle et al. (1995) describe this approach in the context of cross-over designs and point out the disadvantage of loosing some information as the method relies entirely upon within-subjects comparisons.

2.6 Non-Linear Variance Component Models

Nonlinearity is an important theme underlying many current developments in the field of biostatistics and clinical studies. In this section we outline few general aspects of non linearity in connection with components in both regression and no regression situations, that is, with or without the covariates.

Dolby and Freeman (1975) discussed the ML estimation of non linear functional relationships with repeated observations. The analysis for bivariate data was extended to multivariate situations and the error variance was considered to be knows. Previous articles dealing with repeated observations are Villages (1961), Dolby (1972), etc. Dolby (1976) later worked on structural relations of this type. Chan and Mak (1979) assumed

a linear structural relation of the type

$$\left.\begin{array}{rcl} y_i & = & \alpha + \beta X_i \\ y_{ij} & = & y_i + e_{ij} \\ X_{ij} & = & X_i + e_{ij} \end{array}\right\} \begin{array}{l} i = 1, 2, \ldots, n \\ j = 1, 2, \ldots r \end{array} \tag{2.11}$$

with usual assumptions. He found the ML solution to be a root of a fourth degree polynomial. However, it is consistent as the number of replication increases.

We now consider the linear set-up

$$\left.\begin{array}{rcl} y_{ij} & = & \mu + \alpha_i + e_{ij} \\ y & = & X\beta + Zu \end{array}\right\} \begin{array}{l} i = 1, 2, \ldots, n \\ j = 1, 2, \ldots r \end{array} \tag{2.12}$$

The non-linearity may arise from the above formulation in the following ways :

(i) The systematic part $X\beta$ is replayed by non-linear form, as considered by Rudemo el al. (1989) in application to bioassay data.

(ii) The random component α_i and e_i, combine non-linearity. For example, the non-linearity may be approximately modeled as

$$y_{ij} = \mu + \alpha_i + e_{ij} + a_{20}\, \alpha_i^2 + a_{11}\, \alpha_i\, e_{ij} + a_{02}\, e_{ij}^2 \tag{2.13}$$

(iii) The random and systematic parts, in general model, combine non-linearity. One may consider a exponential growth model

$$y_{ij} = exp\, [\mu + (\beta + \alpha_i)\, X_{ij}] + e_{ij}. \tag{2.14}$$

Racine and Poon (1985) adopted such type of approach.

(iv) The essential normal theory based structure may be replaced by an analogous form of the exponential family.

Solomon et at. (1992) discussed in detail the model formulation is a balanced set up. They also proposed the transformation model of the type

$$y_{ij}^{1/\lambda} = \mu + \alpha_i' + e_{ij}', \tag{2.15}$$

the right hand side being a normal theory representation. The transformation models are widely used in practice and often provide a simple basis for structural analysis and interpretation. Choice between the models would naturally depend on the context, as well as on practical considerations.

Soloman also proposed few generalizations of the model and explained the estimates obtained for a medical treatment data. The approximate likelihood function is developed and its accurate performance is examined is examined numerically using examples of exponential regression.

2.7 A Discussion on the ML Method of Estimating Variance Components

The beginning in search for an alternative to ANOVA procedure of estimating VCs appears to lie with Crumpt (1947, 1951). He dealt with the one-way classification for both balanced and unbalanced data and derived equations that have to be solved iteratively. Herbach (1959) derived explicit ML equation for certain balanced models and felt the necessity that such estimators must be non negative. Carbeil and Searle (1976 b) studied a number of these balanced models and obtained the biases and sampling variances of the estimators.

In contrast to the ANOVA approach, the basis requirements of ML estimation is that of assuming an underlying probability distribution for the data. The ML approach to the estimation of VCs has some attractive features. The ML estimators are functions of every sufficient statistics and are consistent. These are also asymptotically normal and efficient. The ML approach is always well defined even for many generalizations of the ANOVA models. Also with ML approach, the non negativity constraints on the VCs or other constraints on the parameter space cause no conceptual difficulties. Moreover, the ML estimates and the information matrix (IM) for a given parameterization of the model can be obtained readily from those for any other parameterization (Harville, 1977).

In late seventies and eighties, many researchers have studies different aspects of ML method of estimating VCs. To name a few : Olsen et al.(1976), Hocking and Kutrer

(1975), Harville (1977).

For the model (2.4), we assume $\underset{\sim}{y} \sim N(\mu\, 1_N, V)$, where $V = Var(\underset{\sim}{y}) = I_a \otimes (\sigma_\alpha^2\, J_{n_i} + \sigma_e^2\, I_{n_i})$ is defined earlier with exception that n_i stands for unbalanced data and $n_i = n$ for balanced data. The log likelihood may be defined as

$$l = -\frac{1}{2}\, N\, ln(2\pi) - \frac{1}{2}(N-a)\, ln\, \sigma_e^2 - \frac{1}{2}\sum_i ln(\sigma_e^2 + n_i\, \sigma_n^2)$$

$$-\frac{1}{2\sigma_e^2}\sum_i \sum_j (y_{ij} - \mu)^2 + \frac{1}{2\sigma_e^2}\sum_i \frac{\sigma_e^2}{\sigma_e^2 + n_i\, \sigma_\alpha^2}\,(y_i - n_i\, \mu)^2 \qquad (2.16)$$

For balanced data, this log likelihood may be written as

$$l = -\frac{1}{2}\, N\, ln(2\pi) - \frac{1}{2}(n-1)a\, ln\, \sigma_e^2 - \frac{1}{2}a\, [ln(\sigma_e^2 + n\, \sigma_n^2)]$$

$$-\frac{1}{2\sigma_e^2}\sum_i \sum_j (y_{ij} - \mu)^2 + \frac{n^2\, \sigma_\alpha^2}{2\sigma_e^2(\sigma_e^2 + n\, \sigma_\alpha^2)}) \sum_i (y_{ij} - \mu)^2$$

$$= -\frac{1}{2}\, N\, ln(2\pi) - \frac{1}{2}a(n-1)\, ln\, \sigma_e^2 - \frac{1}{2}\, a\, ln\, \lambda - \frac{SSE}{2\sigma_e^2} - \frac{SSA}{2\lambda} - \frac{an\,(y_{00} - \mu)^2}{2\lambda}, \qquad (2.17)$$

for $\lambda = \sigma_e^2 + n\, \sigma_\alpha^2$ (Searle et al. (1995). Then for balanced data, the ML estimators may be obtained as

$$\hat{\mu} = y_{00},\ \ MSE = \sigma_e^2,\ \ \hat{\lambda} = \frac{SSA}{a} \qquad \text{and,}$$

$$\hat{\sigma_\alpha^2} = \frac{\hat{\lambda} - \hat{\sigma_e^2}}{n} = \frac{(1 - 1/a)\, MSA - MSE}{n}.$$

Similarly, for unbalanced data, the loglikelihood may be defined as

$$l = \frac{1}{2}\, N\, ln(2\pi) - \frac{1}{2}(N-a)\, ln\, \sigma_e^2 - \frac{1}{2}\sum_i ln\, \lambda_i - \frac{SSE}{2\sigma_e^2} - \frac{1}{2\lambda_i}\sum_i n_i(y_{i0} - \mu)^2, \qquad (2.18)$$

where $\lambda_i = \sigma_e^2 + n_i\, \sigma_\alpha^2$, and the ML estimates may be obtained from the following relations :

$$\hat{\mu} = \sum_i \frac{n_i}{\lambda_i}\, y_{i0} \mid \sum_i n_i,\ \hat{\lambda}_i = \frac{\sum_i y_{i0}/Var(y_{i0})}{\sum_i 1/Var(y_{i0})},$$

$Var(y_{i.}) = \hat{\sigma}_\alpha^2 + \hat{\sigma}_e^2/n_i$;

$$\frac{SSE}{\hat{\sigma}_e^2} - \frac{N-a}{\hat{\sigma}_e^2} + \sum_i \frac{n_i(y_{i0} - \hat{\mu})^2}{\hat{\lambda}_i^2} - \sum_i \frac{1}{\hat{\lambda}_i} = 0$$

and $\sum_i n_i^2(y_{i0} - \hat{\mu})^2/\hat{\lambda}_i^2 = \sum_i \frac{1}{\lambda_i}$.

A detailed discussions on the non-negativity conditions of VCs, the bias and sampling variances are available in Searle et at. (1992).

ML Estimates of Some Linear Models

The ML estimates of four VC models discussed in Section (2.5) are presented in Searle's book (1992).

For general linear model proposed by Laird and Ware, we have,
$Var(y_i) = \sum_i = Z_i\ D\ Z_i' + R_i$ and for $W_i = \overset{-1}{\sum_i}$, we have

$$\left.\begin{aligned} \hat{\alpha} = & = (\sum_i X_i'\ W_i\ X_i)^{-1} \sum_i X_i'\ W_i\ y_i \\ \text{and } \hat{b}_i & = DZ_i'\ W_i(y_i - X_i\ \hat{\alpha}) \end{aligned}\right\} \tag{2.19}$$

The estimate of α maximizes the likelihood based on the marginal distribution of the data and it is also the MVUE. The expression for $\hat{b}_i$ is of course not ML but can be derived by an extensive of Gauss-Markov theorem to cover random effects (Harvill, 1976). The estimate for b_i is also empirical Bayes.

Since $\hat{\alpha}$ and $\hat{b}_i$ linear functions of y, the expression for their S.E. can be easily derived as

$$\begin{aligned} Var\ (\hat{\alpha}) &= (\sum_i X_i'\ W_i\ X_i)^{-1} \\ \text{and} \quad Var(\hat{b}_i) &= DZ_i'[W_i - W_i\ X_i(\sum_i X_i'\ W_i\ X_i)^{-1}\ X_i'\ W_i]\ Z_iD \end{aligned} \tag{2.20}$$

Laird and Ware has also discussed in detail the estimation procedure for unknown variance and the estimation of covariance matrix. They have discussed the use of EM (Expectation Maximization) algorithm for ML estimation of variance components.

The MLE of the implicit parameters in the general linear model can be obtained by maximizing the joint likelihood function

$$l = -\frac{1}{2}\ N\ ln\ (2\pi) - \frac{1}{2} \sum_{i=1}^{n} \left[(y_i - X_i\alpha)'\ \sum_i^{-1}\ (y_i - X_i\ \alpha) + ln\ \mid \Sigma_i \mid\right] \tag{2.21}$$

where N is the total number of observation. The value of α which maximizes the above likelihood for fixed elements of Σ is the GLS estimator $\hat{\alpha}$ defined above. Using the estimate in the of α, the problem reduces to maximize the profile likelihood

$$\hat{l} = \frac{1}{2}[N\ ln\ (2\pi) + \sum_{i=1}^{n}(\hat{e}_i'\ \Sigma_i^{-1}\ \hat{e}_i + ln\ \mid \Sigma_i \mid] \tag{2.22}$$

where $\hat{e}_i = (y_i - X_i\ \hat{\alpha})$.

We note hare that MLE of VCs are biased in small samples. That in why many authors has advocated the use of REML method of estimation.

2.8 REML Method of Estimation VC

A property of ML estimation is that in estimating VCs it does not take into account the degree of freedom that are involved in estimating fixed effects. Although under normality ANOVA estimators are MVUE, ML estimators not do.

The feature of ML not taking account of the degrees of freedom when estimating VCs is overcome by what is known as REML method of estimation. First developed for certain balanced data by Anderson et al. (1952) and Russel el al. (1958), it was extended by Patterson and Thompson (1971, 1974) to mixed model generally.

The basic idea of REML estimation is that of estimating VCs based on residuals calculated after fitting by OLS just the fixed part of the model. REML estimation can also be viewed as maximizing a marginal likelihood.

Let us start with the REML estimation of VC for one-way random model for balanced data. For the one-way ANOVA model under balanced data set up, the likelihood function can be written as

$$L(\mu, \sigma_e^2, \sigma_a^2/y) = (2\pi))^{-\frac{1}{2}an}\sigma_e^{-2[\frac{1}{2}a(n-1)]}\lambda^{-\frac{1}{2}a}\ Exp\left[-\frac{1}{2}\left\{\frac{SSE}{\sigma_e^2} + \frac{SSA}{\lambda} + \frac{y_{00}-\mu)^2}{\lambda/an}\right\}\right] \tag{2.23}$$

Since y_{00} is independent of bath SSE and SSA, the above likelihood can be factored as

$L(\mu, \sigma_e^2, \sigma_\alpha^2/y) = L(\mu/y_{00}\ L(\sigma_e^2, \sigma_\alpha^2/SSE, SSA)$, where

$L(\mu/y_{00} = (2\pi)^{-\frac{1}{2}}\ (\lambda/an)^{-\frac{1}{2}}\ exp[-y_{00} - \mu)^2/(2\lambda/an)]$ and

$$L(\sigma_e^2, \sigma_\alpha^2/SSE, SSA) = (2\pi)^{-\frac{1}{2}(an-1)}\ \sigma_e^{-2}\ \left[\frac{1}{2}\ a(n-1)\right]\ \lambda^{-\frac{1}{2}(a-1)}\ (an)^{-\frac{1}{2}}$$
$$exp\left[\frac{1}{2}\ \left(\frac{SSE}{\sigma_e^2} + \frac{SSA}{\lambda}\right)\right]. \tag{2.24}$$

The last likelihood may be expressed as

$$L(\sigma_e^2, \sigma_\alpha^2/SSE, SSA) = \int L(\mu, \sigma_\alpha^2, \sigma_e^2/y)\ d\mu\ ,$$

showing the marginal likelihood relationship. For 1-way balanced model, this is known as restricted likelihood. The REML equations may be obtained as

$$\begin{aligned} l_{R,\sigma_e^2} &= \frac{-a(n-1)}{2\sigma_e^2} + \frac{SSE}{2\sigma_e^4} \\ \text{and}\quad l_{R,\lambda} &= \frac{-(a-1)}{2\lambda} + \frac{SSA}{2\lambda^2}. \end{aligned}$$

The estimates may be obtained as

$$\begin{aligned} \widehat{\lambda}_R &= SSA/(a-1) = MSA,\ \ \hat{\sigma}_e^2 = SSE/a(n-1) = MSE,\ \text{ and thus} \\ \hat{\sigma}_{\alpha,R}^2 &= \frac{1}{n}\ [MSA - MSE]. \end{aligned}$$

Let us now outline he methodology for the general model. Actually are maximize a reduced log likelihood function obtained by transforming y_i to $\underset{\sim}{Y}_i^*$ where the distribution of $\underset{\sim}{Y}_i^*$ is independent of $\underset{\sim}{\alpha}_i^*$. One such transformation is obtained by taking $\underset{\sim}{Y}_i^* = (I - X(X'X)^{-1}\ X')\underset{\sim}{Y}$, where $\underset{\sim}{Y}^{*1} = \underset{\sim}{Y}_1^{*1} \ldots \underset{\sim}{Y}_n^{*1}]$. The transformation is obviously $\underset{\sim}{Y}_i^* = \underset{\sim}{Y}_i - X_i\ \underset{\sim}{\hat{\alpha}}$ where $\hat{\alpha} = (X'X)^{-1}\ X^1Y$ is simply the OLS estimator of $\underset{\sim}{\alpha}$. It follows that $E(\underset{\sim}{Y}_i^* = 0$ for any $\underset{\sim}{\alpha}$ and in fact distribution of $\underset{\sim}{Y}_i^*$ is independent of $\underset{\sim}{\alpha}_n$. Under this transformation, the reached profile log likelihood can be shown to be

$$\widehat{l}_R = -\frac{1}{2}\left[(N-s)\ ln(2\pi) + \sum_i(\hat{e}_i^1\ \sum_i^{-1}\ \hat{e}_i + ln\mid \sum_i \mid) + ln\mid \sum_i X_i^1\ \sum_i^{-1}\ X_i)\mid\right] \tag{2.25}$$

In both ML and REML, α is obtained by (2.19). The REML estimates of VC are obtained by maximizing the above log-likelihood. Details on REML estimating equations for mixed models is given in the book of Searle et al. (1992). We note hare that sections of REML equations, for all cases of balanced data from mixed model, are same as ANOVA estimators. This result is true whether normality is assumed or not.

For unbalanced data each of ML and REML are to be preferred over any ANOVA method (Searle et al., 1992). This is because the ML and REML estimates are consistent, asymptotically normal, and the sampling dispersion matrix is also known. This provides opportunity to develop confidence interval and testing hypotheses about parameters. If is true that ML and REML estimators are based on normality assumption, but in many situations this assumption is unlikely to be seriously wrong. Of course, the asymptotic variance-covariance property are valid only is large sample sense, and for small samples this may nullify their usefulness. Nevertheless, these properties seem to be sufficiently reliable to have more faith in ML and REML than in the ANOVA method.

Now, to chose between ML or REML — then is no hard and fast answer. Both have the same merits of being based on maximum likelihood principle — and they have the same demerit of computational complexity. ML provides estimators of fixed effects, whereas REML, on its own, does not. But for balanced set up REML solutions are identical to ANOVA estimators which have optimal minimum variance p party. For many researchers this is a strong ground for REML that they prefer it over ML.

2.8.1 The Use of Different Algorithms for Likelihood Estimation

We know that both ML and REML method contain considerable amount of computational complexities. without the aid of high speed computers, obtaining ML and REML estimates of VCs along with the estimates of fixed effects seems to be an impossible task. In this section we review considerations in computing ML and REML estimates and outline algorithms used to the estimates.

In all but the simplest cases, iterative methods must be used to find estimates for the parameters in mixed effects repeated measurements models. The basic iterative methods

Now, to chose between ML or REML — there is no hard and fast answer. Both have the same merits of being based on maximum likelihood principle — and they have the same demerit of computational complexity. ML provides estimators of fixed effects, whereas REML, on its own, does not. But for balanced set up REML solutions are identical to ANOVA estimators which have optimal minimum variance property. For many researchers this is a strong ground for REML that they prefer it over ML.

2.8.1 The Use of Different Algorithms for Likelihood Estimation

We know that both ML and REML method contain considerable amount of computational complexities. without the aid of high speed computers. obtaining ML and REML estimates of VCs along with the estimates of fixed effects seems to be an impossible task. In this section we briefly review considerations in computing ML and REML estimates and outline algorithms used to obtain the estimates.

In all but the simplest cases, iterative methods must be used to find estimates for the parameters in mixed effects repeated measurements models. The basic iterative methods are explicitly based on th derivatives of the log likelihood. These are called gradient methods in the numerical analysis literature. The commonly used method is the Newton-Raphson (NR) method.

Suppose the function is f that we are trying to maximize in the parameter space θ. The iterative formulation would be

$$\theta^{(m+1)} = \theta^{(m)} - \left(H^{(m)}\right)^{-1} \Delta\ f^{(m)} \tag{2.26}$$

where $H^{(m)}$ and $\Delta f^{(m)}$ are Hession and gradient vector respectively. This is NR method. In variance component estimation, the parameter space is $\underset{\sim}{\theta}^{1} = \left[\underset{\sim}{\alpha}^{1}, \underset{\sim}{\sigma}^{2'}\right]$ for ML and $\underset{\sim}{\theta} = \underset{\sim}{\sigma}^{2'}$ for REML. The NR interaction would be (2.26) with $f^{(m)}$ replaced by $l^{(m)}$.

$$\theta^{(m+1)} = \theta^{(m)} - \left(H^{(m)}\right)^{-1} \frac{\partial l}{\partial \theta}\ |_{\theta^{(m)}}$$

With the entries in Hessian H given by

$$\left.\begin{aligned} l_{\alpha} &= X^{1}V^{-1}y - X^{1}V^{-1}X\alpha \\ l_{\alpha\alpha} &= -X^{1}V^{-1}X, \text{ and} \\ l_{\alpha\sigma^{2}} &= -XV^{-1}Z_{i}Z_{i}'V^{-1}(y - X\alpha) \end{aligned}\right\} \quad for\ ML$$

and $\frac{\partial l_R}{\partial \sigma_i^2 \partial \sigma_j^2} = \frac{1}{2}\ tr\left[PZ_j\ Z_j'\ PZ_i\ Z_i'\right] - y'PZ_j\ Z_j'\ PZ_iZ_i'\ Py$. for REML with V as the dispersion matrix of X and $P = K(K'VK)^{-1}\ K'$ under the assumption $K'X = \underset{\sim}{O}$.

Apart from NR same other methods like method of scoring (Jennrich and Sampson, 1986) and Quasi-Newton method (Kennedy and Gentle, 1983) are also available.

There are many techniques that can be applied to reduce computational burden for ML and REML methods. Harville (1977), Jennrich and Sampson (1986) give matrix identities that greatly reduce the size of matrices to be manipulated. Lindstrom and Bates (1988) give a number of details on matrix decomposition that can be exploited to speed up iterations.

An alternative algorithm for calculating ML and REML estimates that differ from NR or scoring method is the Expectation Maximization (EM) algorithm. This method alternates between calculating conditional expected values and maximizing specified likelihoods. This algorithm is due to Dempster, Lair and Rubin (1977). The EM algorithm only generates estimates and does not give variance of the estimates as by product, as do NR and scoring method. To obtain variance estimates extra computation must be performed.

2.9 Some Discussions on the Testing Aspects in Linear Models

There is a long history in the development of F - tests for ANOVA estimators in variance components models. The history of developments is well discussed by Searle (1971). Consider the model

$$y_{ij} = \mu + \alpha_i + e_{ij}. \qquad \begin{aligned} i &= 1, 2, \ldots, a \\ j &= 1, 2, \ldots, n_i. \end{aligned}$$

In the fixed effects model ($\sum_i a_i = 0$). $F = \frac{MSA}{MSE}$ tests the hypothesis H : a_i's are all equal. Under this Hypothesis $F \sim F_{(a-1),a(n-1)}$. In the random effects model, provided the data are balanced, we have

$$SSA/(n\sigma_a^2 + \sigma_\epsilon^2) \sim \chi^2_{(a-1)}, \quad SSE/\sigma_\epsilon^2 \sim \chi^2_{a(n-1)}. \; and \; thus$$

$$\frac{MSA/(n\sigma_a^2 + \sigma_\epsilon^2)}{MSE/\sigma_a^2} \sim F_{(a-1)\,a(n-1)}.$$

which leads to

$$\sigma_\epsilon^2 F/(n\sigma_a^2 + \sigma_\epsilon^2) \sim F_{(a-1)\,a(n-1)} \tag{2.27}$$

Similarly, the tests can be formulated for unbalanced data and the corresponding confidence intervals can be developed.

We note that if the interest is to test the hypothesis $H_0 : \sigma_a^2 \leq \lambda$ against $H_0 : \sigma_a^2 > \lambda$, then the F-test can be formulated on the basis of SSA and SSE. These tests are due to Scheffe (1959).

Weerahandi (1991) developed an alternative testing procedure for VCs in mixed models with generalized p-values. His test was based on minimal sufficient statistics. The p-value for testing $H_0 \cdot \sigma_a^2 \leq \lambda$ is

$$p = 1 - E\left[G\left(\frac{SSA}{n\lambda + SSE/U}\right)\right]$$

where G is the cumulative density function of $\chi^2_{(a-1)}$, and the expectation is taken with respect to $U \sim \chi^2_{a(n-1)}$. Here $U = SSE/\sigma_\epsilon^2$. H_0 can be rejected if the observed value of p is too small, say if $p < 0.05$. Weerahandi also extended the test for 2-way crossed classification models.

2.9.1 Tests Under Unequal Error Variances

It is well known that F-test for 1-way ANOVA is sensitive to the homogeneity of error variances (for example, Brown and Forsythe, 1974, Bishop and Dudewicz, 1978; Tan and Tabatabai, 1986). Many researchers discussed the effect on type 1 error and conclude

that the effect is not serious when the group sizes are equal. The effect of heterogeneity on the F - test with respect to both size and power is described by Krutchkaff (1988). He discussed the drawbacks of well-known Kruskal-Wallis Test which is often recommended in situation where the errors are heterogeneous. Krutchkaff proposed an alternative, called K-Statistic which is developed on the principle of weighting each component with its information (one over its estimated variance). His average for each group is

$$\bar{y}_i = \frac{\sum_i y_{ij}/\hat{\sigma}_i^2}{\sum_i \frac{1}{\sigma_i^2}}.$$

where y_{ij} represents the j^{th} observation on i^{th} group, and $\hat{\sigma}_i^2$ is an estimate of the variance of y_{ij}. Using these the overall average becomes

$$\bar{y} = \left(\sum_i \bar{y}_i n_i/\hat{\sigma}_i^2 \right) / \sum_i n_i/\hat{\sigma}_i^2.$$

The test statistic that can be used to test the equality of group means is given by

$$K = \sum_i \left[(\bar{y}_i - \bar{y})^2 \; n_i/\hat{\sigma}_i^2 \right] /(k-1).$$

If pooled estimate of variance is used for each $\hat{\sigma}_i^2$ then this K-statistic is identical to the F-statistic. The distribution of K and its relation to likelihood ratio (LR) statistic under heterogeneity is studied by Good (1986). Krutchkoff (1985) obtained the critical values for K-statistic.

Weerahandi (1991), Zhou and Mathew (1994) used p-values to construct exact procedures for comparing VCs in the mixed models. By taking a generalized approach to finding p-values, the classical F-test of the 1-way ANOVA is extended to the case of unequal error variances by Weerahandi (1995). In the context of regression, Koschat and Weerahandi (1992) provided a class of tests based on p-values to compare parameters of regression models in the presence and absence of common parameter.

2.9.2 Testing For General Linear Model

We now discuss the inference problem in general linear mixed effects model. The large sample tests of hypotheses can be carried out using the Wald Chi-square test defined as

$$T^2\left(\widehat{\theta}\right) = \widehat{\theta}'\, C'\, [C\, \widehat{\Omega}\, C'\,]^{-1}\, C\, \widehat{\theta} \sim \chi^2_{rank(C)}.$$

The alternative tests include the LR test

$$2L\left(\widehat{\alpha}_{full}, \widehat{\theta}_{full}\right) - 2L\left(\widehat{\alpha}_{reduced}, \widehat{\theta}_{reduced}\right) \sim \chi^2_{df(full)-df(reduced)}$$

As an improvement on the Wald test in small samples, one may also use the MANOVA like F-test

$$F(\widehat{\alpha}) = (2\, s_o n_o - s_o + 2)(s_o C u)^{-1}(n-q)^{-1}(n-q)^{-1}\, T^2(\widehat{\alpha})$$

for testing the hypothesis $H_0 : (U' \otimes C)\, \alpha = \underset{\sim}{0}$. Here C is a $(c \times q)$ contrast matrix of within-subject comparison. The critical values for this approximate test are obtained from tabled values of F-distribution : $F_{(cu,\ 2s_o n_o - s_o - 2)}$. Here, $c = rank(C)$, $u = rank(U)$, $s_o = min(c, u)$, $n_o = [(n-q) - u]/2$. This small sample test is asymptotically valid as discussed by Vonesh and Chinchilli (1997).

Under normality assumptions, the goodness of fit and model selection for repeated measures models can be carried out using LR test for nested models and Akaike's Information Criteria (AIC) or Schwarez's Bayesion Criteria (SIZE) for non-nested models. Let us consider the following two linear models (nested).

$$\text{Model I} : y_i = [X_{i1}\ X_{i2}] \begin{bmatrix} \underset{\sim}{\alpha}_1 \\ \underset{\sim}{\alpha}_2 \end{bmatrix} + Z_i\, b_i + e_i$$

$$\text{Model II} : y_i = X_{i1}\, \alpha_1 + Z_i b_i + e_i$$

To test $H_O : \underset{\sim 2}{\alpha} = 0$ where $\underset{\sim 1}{\alpha}$ has dimension $(s_1 \times 1)$ and $\underset{\sim 2}{\alpha}$ has dimension $(s_2 \times 1)$, the LR test is

$$2l_1(\hat{\underset{\sim 1}{\alpha}}, \hat{\underset{\sim 2}{\alpha}}, \theta) - 2l_2(\hat{\underset{\sim 2}{\alpha}}, \theta) \sim \chi^2_{s_2}.$$

where l_1 & l_2 are log-likelihoods for full and reduced models.

The LR test for nested model is some what cumbersome particularly if one wishes to examine the robustness of models to changes in mean and variance-covariance. An alternative approach is due to Akaike (1974) which uses likelihood based measure with adjustment of parameters in the model. The AIC is defined by

$$\begin{aligned} AIC &= \hat{l}(\hat{\theta}) - s^* \\ &= \hat{l}_R(\hat{\theta}) - s^* \end{aligned}$$

where $\hat{l}$ and $\hat{l}_R$ are profile and restricted profile log-likelihoods evaluated at $\hat{\alpha}(\hat{\theta})$ and $\hat{\theta}$. $s^* = dim(\alpha) + dim(\theta)$. The alternative approach SBC is defined as

$$\begin{aligned} SBC &= \hat{l}(\hat{\theta}) - s^* \, ln(N)/2 \\ &= \hat{l}_R(\hat{\theta}) - s^* \, ln(N - s)/2 \end{aligned}$$

where N is the total number of observation. While comparing models, larger the values of AIC or SBC, the better the fit.

Alternative criteria similar to R^2-type measure may also be used to assess the fit, as proposed by Kralseth (1985). Suppose our model is the hypothetical model and Model II be the null model. Also suppose for Model I. $var\ (e_i) = \sum_{i1}$, and $var\ (e_i) = \sum_{i2}$ for Model II. Let $\hat{y}_{i1}$ be the fitted value for Model I and $\hat{y}_{i2}$ be that for 2nd one. Then R^2 measure is

$$R_i^2 = 1 - \frac{(y_i - \hat{y}_{i1})' \, V_i^{-1} \, (y_i - \hat{y}_{i1})}{(y_i - \hat{y}_{i2})' \, V_i^{-1} \, (y_i - \hat{y}_{i2})}$$

where V_i is any positive - definite matrix Either Σ_{i1} or Σ_{i2} may be chosen for V_i. Since null model and Σ_{i2} remain fixed, choosing $V_i = \Sigma_{i2}$ would be consistent with desire of

having a goodness-of-fit measure that can be compared across different hypothesized model.

Using the criteria established by Kralseth. Vonesh et. al. (1996) define a model concordance correlation coefficient that can be used as an alternative to R^2 defined above. It is defined as

$$r_c = 1 - \frac{\sum_i (y_i - \hat{y}_{i1})' (y_i - \hat{y}_{i1})}{\sum_i (y_i - \bar{y}\, j_{p_i})' (y_i - \bar{y} j_{p_i}) + \sum_i (\hat{y}_i - \hat{y}\, j_{p_i})' (\hat{y}_i - \hat{y} j_{p_i}) + N(\bar{y} - \hat{y})^2}$$

where j_{p_i} is $(p_i \times 1)$ unit vector, $\bar{y} = \sum y_{ij}/N$. $\hat{y} = \sum \hat{y}_{ij}/N$ are grand means y_{ij} and $\hat{y}_{ij}$ respectively.

Both R^2 and r_c provide measure of goodness-of-fit. However, r_c may be better in that it is directly interpretable as a concordance correlation between observed and predicted values. As a measure of agreement, its value reflects how well a scatter plot of y_{ij} versus $\hat{y}_{ij}$ falls about the line of identity. Thus r_c does not require specification of a null model since the line of identity serves as a point of reference. The range of r_c is $[-1, 1]$. We note here that r_c can easily be modified to allow for heterogeneity in $\underset{\sim}{y}_i$ by basing on certain transformation for $\underset{\sim}{y}_i$.

CHAPTER - 3

ASSOCIATION OF BIOCHEMICAL PARAMETERS IN CTC TEAS WITH TASTER'S CHOICE

3.1 Introduction

In the chapter we try to associate the chemical parameters in tea with the Tea Taster's choice, so that the significance of different biochemical quality parameters in explaining different quality attributes can be statistically assessed. We note here that the tea quality attributes are not directly measurable and are evaluated by the Tea Tasters using their sensory methods. In regression setup, the ordinal scores given by a Taster, on a particular attribute, represent the response (dependent) variable. We restrict our study to those data sets, where only a single Taster evaluates the tea samples in terms quality attribute(s). We note here that the general practice in Tocklai Experimental Station, India, is to get the sensory analysis done by a single Taster only, possibly due to cost consideration.

Only a few statistical studies have been made so far in these lines (McDwell et al., 1991). Most of the studies are based on only total correlation between biochemical information and Taster's scores on individual liquor characteristic. Linear equations were set up without proper data diagnostics.

The series of papers by Biswas et al. (1971) attempt to associate Taster's scores on attribute like strength, briskness, brightness etc. with the overall quality scores. Associations have also been investigated between biochemical quality parameters and Taster's choices on various liquor characteristics and on cash valuations; and thereafter the biochemical and botanical implications of the results have been studied. Their study

was based on the North-East Indian plain black CTC and Orthodox teas. According to their findings, regardless of Tasters and method of manufacturing, four biochemical constituents : total oxygen uptake of unprocessed tea shoots, TF, theogallin (TG) and epicatechin gallate (ECG) of black tea, are the main guiding constituents of N-E Indian plains tea. For the orthodox tea samples, the enzyme activity and ECG of unprocessed tea shoots along with water soluble solids (WSS) are of utmost importance.

Although Biswas et al. (1971) contributed to a great extent in understanding the association of different chemical quality parameters with the Taster's choices, they could not provide sufficient information on the behaviour of different biochemical quality parameters towards individual quality attributes, though they had a strong data base for these studies. Another important aspect left untouched is the subjectivity of Taster's choice while formulating the association. However, their study helped the biochemists and agronomists associated with the tea research to a great extent.

Some progress have been made in relating certain groups of tea constituents to quality. Successful relationships have been demonstrated between the total theaflavins levels of Central African black teas and sensory evaluations or prices (Cloughley, 1981 and Ellis and Cloughley, 1981). Such relationships were positive but less successful for Kenya black teas (Owuor et al., 1986). The success obtained in the regression between prices and total (Flavognost) theaflavins for Central African black teas led to the suggestion that total TF level is the objective quality parameter (Davis, 1983) which may be used as standard in black teas. But this suggestion was opposed by producers whose total TF levels show little relation with sensory evaluations (Othieno and Owuor, 1984) and it was argued that there may be other more important black tea quality parameters. Indeed, some Kenyan black teas subsequently showed better relationship between the aroma and sensory evaluation (Owuor, 1992). Unfortunately, we can not site a single study in this line specific to Indian black teas (except those of Biswas et al., 1971).

McDwell et al. (1991) studied the black tea sample using HPLC, collected from seven countries. Principal Component method was used to highlight the characteristic differences in phenolic constituent levels for different countries. Linear regression technique was

used to investigate the relationships between price (score) and phenolic constituent levels. They supported the great importance of phenolic chemistry in the determination of tea quality. The similar phenolics (TFs) appeared to be capable of explaining intra-regional as well as inter-regional variation in quality or value. They took the Tasters' scores on structured scale but did not address the problem of subjectivity in Taster's choices.

From the studies made so far, it is clear that the quality/value depend on a complex of biochemical constituents and are region dependent. That is why it is more important to study a wider range of biochemical quality parameters and their influence on the overall quality or value. We try to study this aspect on the basis of the data provided by the tea industry. The insufficiency in data base limits our study, as information only on a few chemical parameters (that too for CTC teas only) have been provided to us. We cannot claim completeness of our statistical analysis which could otherwise provide a great statistical support service to the tea industry so far the aspects of tea quality assessment is concerned.

3.2 Multiple Regression Analysis With Measurement Error in Response : A Discussion

When a single Taster's score corresponding to a set of chemical information (explanatory variables) is available for a particular tea sample, we may formulate the linear regression model as

$$y_i + u_i = x_i'\beta + e_i$$

where u_i represents the error due to Taster (i.e., corresponding to the observed value of the response) and e_i is the random error component which includes the effect of unobserved or unobservable chemical components in explaining the quality. The above formulation typically represents a regression model with error-in-variable (EIV) in the regrassand.

The effects of measurement error in the explanatory variables have long been recognized (Fuller, 1987; Pal, 1981). But error in the response variable is often ignored. A reason is that errors in the response in standard linear model are inconsequential so long

as the asymptotic properties of the conventional least square (LS) estimators are concerned. The usual LS method gives unbiased and consistent estimates of β coefficients. With response measurement errors, the response errors get confounded with the equation error, and the effect is to increase the standard errors of the β estimates. Thus, response errors can be ignored if we are only interested in the estimates of regression coefficients, especially in the simple situations. However, in more complicated regression models, certain types of response errors can not be ignored and it is important to explicitly account for the response error in regression analysis. A good text in this line is due to Carrol et al. (1995).

Consider our situation where the measurement error u_i is additive and is independent of e_i (assumption). In this case the response measurement error can be ignored if the regression variance is homoscedastic. Here the only effect of measurement error is that the MSE is $\sigma_*^2 = \sigma_e^2 + \sigma_u^2$ and not σ_e^2. Thus, if we are not interested in estimating separate VCs (σ_e^2 *and* σ_u^2), then the response error can be ignored. Even in case of heteroscedastic situation, the response error can be ignored under certain conditions. We note here that the variance components of the model are not identifiable without repeated observations on the response variable.

3.3 Analysis of 4 Sets of Autumn Flush Data

The four sets of CTC tea data (Data Set 1) have been introduced in the introductory chapter. The biochemical parameters measured are TF, TR, B, TLC and total solvable solids (TSS). We note here that the information on Caffeine is not available which is otherwise known to be a very important biochemical parameter so far the CTC black tea is concerned. The same Taster evaluated the CTC tea samples in terms 'quality' and 'value' on a structured scale of 0-10 points. The four sets of CTC samples were collected from the Tocklai experimental garden over four years in the autumn flush period only. The manufacturing/drying process is the same for all the samples.

The basic statistics for the four sets of samples are presented in Chapter 1. The range

of correlation coefficients for all the variables under study are presented in Table 3.1. The mean profiles for all the variables may be studied. We may test if the profiles for means of the variables for the four data sets under study are parallel and coincide. Here we are interested to test whether the means of variables under study differ significantly for all sets of samples or not. Similar test may be performed for the standard deviations as well. If both means and standard deviations do not differ significantly, we may pool the four data sets together and study the association. As discussed in the second chapter, the tests for parallel and coincident profiles are the F tests based on T^2 statistics. The profile analysis may be performed using the SAS or STATISTICA computational packages.

For the four sets of data we have $n_1 = 25, n_2 = 23, n_3 = 25$ and $n_4 = 21$. For the mean profiles the estimated value of T^2 is 3.67 and the 5% critical value of F with 1 and 90 degrees of freedom is 3.92. Thus we may accept the hypothesis of parallel and coincident mean profiles at 5% level of significance. But the same hypothesis does not hold for the profiles of standard deviations, as the calculated value of T^2 is 12.89. It is clear from the Table 1.1 of Chapter 1 that the values of SD differ widely for most of the variables under study.

The most surprising aspect in the data is the variation in the correlation coefficients between different chemical parameters. As the four sets of CTC samples have the same origin and received the same manufacturing process, it is expected that the correlation pattern between the biochemical quality parameters would be similar. But as may be observed from Table 3.1, the ranges of correlation coefficients are wide enough. The widest among them is the correlation coefficients between B and TR (-0.44, 0.60). This is really surprising. Also, the correlation between chemical parameters (TLC and TSS) and Taster's scores on Q and V are not significantly high. Though the sensory scores are given by the same Taster and the average levels of TF and TR do not differ significantly for all sets of samples, even than the correlation coefficients between quality attributes and TF and TR vary widely. This is partially due to the fact that the Taster's choice is subjective.

For all four sets of samples the 'heteroscedasticity' problem has been the common feature. We have used the χ^2 test for $\hat{e}^2$ and $\hat{y}$ and the $\hat{e}^2$ on X (Brusch-Pagan test, 1980). To check for the 'multicollinearity' we have used the conditional index (CI) test, which is defined as the square root of the ratio of maximum and minimum eigen values (Judge et al., 1985). If this ratio lies between 10 and 30, there is evidence of moderate to severe multicollinearity. For all the four sets of samples under study, these ratios have been found to be much below 10. In the presence of multicollinearity, we could opt for the Principal Component (PC) regression or the Ridge Regression technique (Judge et al., 1985).

We have associated all the five chemical parameters with Q and V respectively. Owing to the problem of heteroscedasticity, we fit the linear regression model with **dependent variable heteroscedasticity**. For the model $y_i = x_i'\beta + e_i$, we assume that e_i is a zero mean, serially uncorrelated process with variance function h_i. A survey of approaches to the specification of h_i is available in Judge et al. (1985). The dependent variable heteroscedasticity form applied in our study is $h_i = (x_i'\beta)^2\alpha^2$, where α is a scalar parameter. This may be easily applied in SHAZAM econometric package which calculates the parameters using Quasi-Newton method. The regression results for all sets of CTC data

are presented in Table 3.2 to Table 3.5. In the association of chemical parameters with Q, the estimated regression coefficients of TF are positive and statistically significant at 5% level only in two cases. The slope coefficients associated with TR are all significant except for the fourth set of samples. The estimated coefficients for B are negative for three sets of samples and are significant only in two cases. The significant influence of TLC on Q is observed only for the second data set. TSS is significant only for the fourth data. The sum of squares of errors (SSE) values are fairly small in all the cases. The values of adjusted R^2 ranges from 0.36 to 0.54.

The association of chemical parameters with V reveals the significant influence of overall TF level for the first three sets of samples. TR is significant only for the first set. B is significant only in two cases. TLC and TSS are insignificant in all cases. Comparing the values of $\bar{R}^2$, SSE and loglikelihood ($ln\ L$), we may say that the fit is better with V than that with Q for all the four samples.

The TF - TR interaction

We tried the regression with TF and TR only. On the suggestion of the biochemists, we incorporate the $TF \times TR$ interaction effect in the model. Here the interaction variable tend to be correlated rather strongly with the individual TF and TR. However, in case of only two variables the problem would not be that serious as it is likely to be with large number of interaction and higher order components. In simple cases the difficulties can be avoided by orthogonalizing the product and power terms with respect to the predictors from which they are constructed (Aiken and West, 1991). A simple procedure of orthogonalization may be to represent the interaction variable by the residual part. The procedure may be described as follows:

We may regress $TF \times TR$ on TF and TR (fit the model $TF \times TR = a + b_1\ TF + b_2\ TR +\ residual$), and save the residual as new variable TF*TR. Note that TF*TR has zero mean and correlates zero with both TF and TR, because it is a residual. With this new variable (TF*TR), the regression model would be

$$Quality\ attribute = a + b_1 TF + b_2 TR + b_3 TF * TR + error.$$

Here we may use the standard t-tests for the significance of regression coefficients.

For all the four data sets, the TF*TR interaction component has been insignificant. This is certainly not supporting the general perception among the chemists about the importance of studying the interaction effects of these compounds. We note here that we have studied only the total TF and TR levels. May be, different TFs and TRs would guide the interaction effects and their consequences towards other volatile compounds like caffeine. That is why it is necessary to study the HPLC data on different TF and TR levels.

From the regression analysis on available data, we may say that TF is the most important chemical parameter in explaining the valuation of CTC tea. We may observe the variability in the partial correlations presented for all the chemical parameters. Similar are the findings of Biswas et al. (1971) for the N-E Indian plain black teas and Ellis and Cloughley (1981) for Central African Black teas. However, detailed study is necessary with several other important chemical parameters. Different TFs and TRs need to be studied separately, specific to the flush periods and on a continuous basis. The relationship between different chemical parameters and the pattern of their variation over samples and also over the flush periods are some important technical aspects which needs to be addressed in detail while associating the biochemical parameters with sensory analysis.

3.3.1 Analysis of Data of Set 3

As mentioned in the first chapter, this set of data was provided by the Tata Tea Ltd. (India). For a set of 50 CTC tea samples, the biochemical parameters measured are TF, TR, TLC and B. Analysis of TF is done by the method of Ullah (1984), based on the liquid-liquid extraction in presence of $Na_2\ HPO_4$. TLC is measured from the whole aqueous extract being diluted with methanol. Duplicate analysis of each were done and the brightness is calculated from the measured biochemical parameters. The Taster's score is given on a 0-5 point scale. There are only three distinct values in the scores, viz., 2.6, 3.0 and 3.4.

The correlation matrix and other basic statistics for this set of samples has already

been presented in the introductory chapter.

Due to the multicollinearity problem (on the basis of CI test) we opt for the Ridge Regression technique here. We have also tested for the heteroscedasticity, and no evidence is found against the hypothesis of homoscedasticity. The estimates obtained using Ridge Regression is compared with those obtained using the Ordinary Least Square (OLS) method.

The OLS fit in this data is very poor and all the coefficient estimates came out insignificant with relatively higher value of R^2. This actually indicates the presence of possible multicollinearity. The OLS assumes normality of residuals. But OLS residuals for this data do not support the assumption of approximate normality. But the residuals obtained using Ridge Regression technique follow approximate normality with a slightly flat right-hand tail. As an alternative, we may also think of a robust estimation procedure, like the Least Absolute Deviation (LAD) technique. This technique is reasonably efficient irrespective of the form of the error distribution and is elaborately discussed by Judge et al. (1985). But for the given samples, the LAD technique do not give a better fit in comparison to the Ridge Regression. We do not present the estimates obtained using LAD technique here. The OLS and Ridge Regression results are presented in Table 3.6.

We note here that of the 50 samples, only of 4 received the highest score (3.4). The medium category quality rating (3) is received by 30 samples and the remaining 16 samples are evaluated as poor or inferior score (2.6). We may divide the data into two categories on the basis of Taster's choice: The first group contains all those samples which received the scores 3.4 and 3. The second group includes 16 samples that received the poor score. We may test if the chemical parameters differ in their average level over these two groups. Tests reveal that only the average level of TR differ significantly at 5% level (t- test). We may also introduce dummy variables to test the significance of difference between estimated regression coefficients for the chemical parameters. However, for the given set of data only the intercept varies significantly as evident from the estimated value of dummy coefficient ($\hat{\beta}_{dummy}$), which is positive for both OLS and ridge regression. In the linear regression model, the dummy variable is of the form: $d_i = 0$ if i^{th} sample receives

lower rating and $d_i = 1$ otherwise.

We note here that regression estimates are obtained for all samples and separately for first 40 and last 40 samples. This is done to predict the first and last 10 samples of the data set (Paul, 1999). Prediction performances using both Ridge regression and LAD method are presented in Table 3.7. As may be observed from the Table, the prediction performance is better using the Ridge Regression method.

As may be observed from Table 3.6, all the biochemical parameters have came out to be significant at 5% level along with the intercept. The coefficient estimate for B is negative. For the comparison of fit, the values of $\bar{R}^2$, SSE, loglikelihood, Akaike Information Criteria (AIC, 1973) and Final Prediction Error (FPE) are considered. Lower the values of AIC and FPE, better is the fit when we compare different nested or non-nested models. These are likelihood based criteria and remain same (theoretically) when we compare the estimates using OLS and Ridge techniques.

3.4 Multiple Response Regression Model

The Tasters may evaluate the tea samples in terms of different quality attributes (e.g., strength, quality, etc.) on a structured scale. The scales may differ among the attributes as well as among the assessors. Dealing with scores on different scales would be more complicated. We consider the situation where a single Taster evaluate the tea samples in terms of different quality attributes using the same structured scale, as is the situation in Data Set 2. This is a multiple response situation where for a given set of regressors (biochemical measurements), there are more than one response variables.

In regression set up, the approach of simultaneous equations systems for multiple response data is appropriate and standard techniques are well developed in the literature. As a special case, when the reduced form equations are same as the structural form equations as in the Zellner's seemingly unrelated regression equations (SURE), the system is always identified and the system equations may be estimated using the techniques discussed in Zellner (1962). We may also assume specific correlation structure of the disturbance terms. Such a specification is reasonable when estimating a number of related functions and the error components for these functions are likely to exhibit some correlation. This correlation between different disturbances at a given point of measurement is known as contemporaneous correlation. The disturbances at a given point of measurement might be expected to reflect some common unmeasurable or omitted influences. Contemporaneous correlation could be the result of these common factors that are not included in the regression.

We outline the multiple response regression model and the estimation procedure in the following subsection.

3.4.1 Model and Estimation

Consider the following generic situation. For i^{th} response variable, the linear regression set up for a set of n samples is given by

$$\underset{\sim}{y}_i = X_i \underset{\sim}{\beta}_i + \underset{\sim}{e}_i, \quad i = 1, 2, \ldots p \tag{3.1}$$

where $\underset{\sim}{y}_i$ and $\underset{\sim}{e}_i$ are $(n \times 1)$ vectors, X_i is $(n \times k)$ matrix of known regressors and $\underset{\sim}{\beta}_i$ is $(k_i \times 1)$ vector of unknown coefficients. Stacking the data for all p response variables, we may write the multiple response regression model as

$$\underset{\sim}{y} = X \underset{\sim}{\beta} + \underset{\sim}{e} \tag{3.2}$$

where $\underset{\sim}{y} = (\underset{\sim}{y}_1, \underset{\sim}{y}_2, \ldots, \underset{\sim}{y}_p)'$, $X = diag\ (X_1, X_2, \ldots, X_p)$, $\underset{\sim}{\beta} = (\underset{\sim}{\beta}_1, \underset{\sim}{\beta}_2 \ldots \underset{\sim}{\beta}_k)'$, and $\underset{\sim}{e} = (\underset{\sim}{e}_1, \underset{\sim}{e}_2 \ldots \underset{\sim}{e}_p)'$. Here $\underset{\sim}{y}$ and $\underset{\sim}{e}$ has the dimension $(np \times 1)$, X is $(np \times k)$ and $\underset{\sim}{\beta}$ is $(k \times 1)$ with $k = \sum_i k_i$.

Suppose e_{ij} be the error component for j^{th} sample in i^{th} equation ($j = 1, 2, \ldots n$). Introducing contemporaneous disturbance correlation, but no correlation between samples, we may write $E(e_{ij}, e_{lk}) = \sigma_{il}$ for $j = k$ and 0 for $j \neq k$. It implies $E(\underset{\sim i}{e}\,\underset{\sim l}{e'}) = \sigma_{il} I_n$ for model (3.1) and the covariance matrix of the error component in model (3.2) would be

$$E(\underset{\sim}{e}\,\underset{\sim}{e}') = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \ldots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \ldots & \sigma_{2p} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \sigma_{p1} & \sigma_{p2} & \ldots & \sigma_{pp} \end{bmatrix} \quad \otimes \quad I_n = \Sigma \otimes I_n = \Phi.$$

Dielman (1989) discusses the nature of gain when viewing the equations as a system rather than estimating them one at a time by OLS method. The GLS estimator of $\underset{\sim}{\beta}$ may be obtained as $\underset{\sim}{\hat{\beta}} = (X'\,\Phi^{-1}\,X)^{-1}x'\,\Phi^{-1}y$ by minimizing $\underset{\sim}{e}'\,\Phi^{-1}\,\underset{\sim}{e} = (\underset{\sim}{y} - X\underset{\sim}{\beta})'\Phi^{P-1}(\underset{\sim}{y} - X\underset{\sim}{\beta})$. The estimator of $\underset{\sim}{\beta}$ is unbiased with covariance matrix $[X'\Phi^{-1}X]^{-1}$.

Judge et al. (1988) discuss the estimated generalized least squares (EGLS) method of estimating the β for unknown $\sum$. Here we use the LS residuals $\underset{\sim i}{\hat{e}} = \underset{\sim i}{y} - X_i\,\underset{\sim i}{b}$ to obtain the estimator $\widehat{\sum}$. $\underset{\sim i}{b}$ is the OLS estimator of $\underset{\sim}{\beta}$. It may be shown that the ij^{th} element in $\widehat{\sum}$ is of the form

$$\hat{\sigma}_{ij} = \frac{2}{n}\underset{\sim i}{\hat{e}}'\underset{\sim j}{\hat{e}}, \quad i, j = 1, 2, \ldots p. \tag{3.3}$$

It is easy to obtain the EGLS estimator of regression coefficient as $\underset{\sim SUR}{\hat{\beta}} = [X'(\widehat{\sum}^{-1} \otimes I_n)X]^{-1}\,(\widehat{\sum}^{-1} \otimes I_n)y$. This is the Zellner's SUR estimator.

We may use an iterative approach to obtain the SUR estimator. After obtaining $\hat{\beta}_{SUR}$. We recomputed the residuals and apply the formula (3.3) to get revised estimate of σ_{ij}. We substitute these revised estimate into $\widehat{\sum}$ and recomputed the coefficient estimates. Convergence criteria may be defined to achieve locally efficient estimates. Hiller et al. (1986) has shown that the iterative estimator and the ML estimator are equivalent. It is known that EGLS estimator is asymptotically superior to LS counterpart. The small sample properties of SURE have been studied by many researchers. We may simultane-

ously estimate $\underset{\sim}{\beta}$ and $\sum$ using likelihood based procedures. The asymptotic properties of ML estimators would be identical to the EGLS estimators of $\underset{\sim}{\beta}$. But Monte Carlo evidence suggests that the iterative ML estimator is not uniformly better than the EGLS estimator in case of finite samples (Judge et al. 1988).

We now consider the problem of introducing an error component with the Taster's choice. For the i^{th} response variable, the regression model may be written as

$$\underset{\sim i}{y} + \underset{\sim i}{u} = X_i \underset{\sim i}{\beta} + \underset{\sim i}{e}, \quad i = 1, 2, \ldots, p,$$

where $\underset{\sim i}{u} = (u_1, u_2, \ldots, u_j)'$ is the vector of response error component for the i^{th} response variable. $\underset{\sim i}{u}$ and $\underset{\sim i}{e}$ are assumed to be independent with each elements in $\underset{\sim i}{u}$ and $\underset{\sim i}{e}$ follows $(0, \sigma^2_{ui})$ and $(0, \sigma^2_{ei})$ respectively. Here the only effect of response measurement error is that the MSE is $\sigma^2_* = \sigma^2_{ui} + \sigma^2_{ei}$ and not σ^2_{ei}. Thus the residual mean square is increased only without any effect on coefficient estimates. This is as well true for the multiple response model. So, the problem of response measurement error is otherwise harmless, so far the estimation of regression coefficients are concerned. Anyway, we can not help the estimation procedure without repeated observations on each response variable, if separate variance components for the measurement errors are of special interest. Had repeated measurements on each response variable been available, we could formulate the multiple response error component regression model.

3.4.2 Analysis of Data Set 2

The Data Set 2 has already been introduced in the first chapter. The three response variables on the basis of Taster's sensory scores are 'quality' (Q), 'strength' (S) and 'total valuation' or 'overall quality' (V). All the quality attributes have been evaluated on a same structured scale. On the basis of the biochemical knowledge discussed in the first chapter and the perception of the Tea Taster of Tocklai Experimental Station, we relate Q with TF, TR, BR and TSS; V with BR, TLC and S. It is believed that Q and V are related, as V is evaluated considering all the higher characteristics together. We have the

following systems of equations.

$$\begin{aligned} Q &= \beta_0 + \beta_1 TF + \beta_2 TR + \beta_3 BR + \beta_4 TSS + e_1 \\ V &= \beta_0' + \beta_1' BR + \beta_2' TC + \beta_3' S + e_2 \end{aligned}$$

The iterative EGLS method of estimation is applied to obtain parameter estimates along with their standard errors. The estimates obtained applying SUR procedure along with different test criteria are presented in Table 3.9. For estimation the convergence tolerance was set at 10^5 and local convergence was achieved after only 9 iterations, obviously showing very low estimation cost.

We have applied both Breusch-Pagan lagrange multiplier (B-P LM test, 1980) and the likelihood ratio (LR) test of Conniffe (1982a) to investigate whether the estimated covariance matrix, $\hat{\Phi}$, is diagonal or not. The null hypotheses are of no correlation among response variates. We note here that both LM and LR test statistics have, asymptotically, chi-square distribution with 1 df. The calculated values of χ^2 for LM and LR tests are 15.827 and 29.552 (Table 3.9). Clearly the hypothesis of diagonality of estimated covariance matrix is rejected here with a strong evidence of the presence of contemporaneous correlation. Hence the application of SUR method is justified here. An improvement in the precision of the estimates may be observed applying SUR method against the OLS counterpart (Table 3.8 and Table 3.9). Note that S.E. of the regression coefficients from EGLS are smaller than those from OLS, resulting from the gain in efficiency due to EGLS over OLS. The OLS fit of the equation involving Q shows very poor performance with low R^2 value and very high residual sum of squares (SSE). Similar is the performance of OLS fit for the equation involving V.

From Table 3.9 we observe that the biochemical parameters TF, TR, BR and the quality attribute S are statistically significant at 5% level of significance. BR is included in both the equations. It may be observed that BR is statistically significant to explain quality (Q) but not that for V. In this situation we apply both asymptotic normal test and the Wald's chi-square test to investigate the hypothesis of no difference between the estimated coefficients $\hat{\beta}(BR)$ and $\hat{\beta}'(BR)$. The calculated values of normal statistic and

the χ^2 statistic (with 1 d.f.) are -1.7272 and 2.9832 respectively with p-values 0.0841 in both cases; clearly indicating no significant difference between the estimated coefficients. We note that for large samples these test statistics are equivalent. We may also test whether all the regression coefficients in SUR model are zero or not. The test statistics discussed by Dielman (1989), is a chi-square one which following central χ^2 distribution with $(p-1)k$ d.f. Here p denotes the number of equations and k denotes the number of regression parameters. The estimated value of χ^2 (139.05, presented in Table 3.9) strongly suggests non-zero coefficients. The value of system R^2 (0.78) is also reasonably high.

Table 3.1 : Range of correlation coefficients between the variable for four sets of samples

Q	1
V	(0.92 - 0.99)
TF	(.20 - .52) (.19 - .56)
TR	(.30 - .55) (.29 - .55) (.16 - .56)
B	(- .33 - .64) (- .29 - .07) (.13 - 0.7) (- .44 -.6)
TLC	(- .22 - .46) (- .20 - .39) (.13 - .46) (- .32, .14) (.20 - .44)
TSS	(- .05 - .44) (- .04 - .26) (- .01 - .52) (- .26, .60) (.17 , .56) (.05, .37)

Table 3.2 : Regression results for first set of samples

	Estimates	S.E.	t-ratio	Partial Correlation
$\hat{\beta}_0$	-35.759	16.77	2.132	-0.45
$\hat{\beta}_{TF}$	18.519	5.195	3.564	0.64
$\hat{\beta}_{TR}$	0.6595	0.2050	3.224	0.61
$\hat{\beta}_B$	-0.5245	0.1308	4.01	-0.69
$\hat{\beta}_{TLC}$	1.1435	.9464	1.208	0.27
$\hat{\beta}_{TSS}$	-0.6461	0.439	1.472	0.33
$\bar{R}^2$	0.54			
SSE	0.0734			
$ln\ L$	-39.81			
$\hat{\beta}_0$	-2.0231	17.08	0.1184	-0.03
$\hat{\beta}_{TF}$	5.2913	2.198	2.4073	0.63
$\hat{\beta}_{TR}$	0.2660	0.1200	2.216	0.46
$\hat{\beta}_B$	-.1752	0.0782	2.240	-0.47
$\hat{\beta}_{TLC}$	0.2866	0.5132	0.5595	0.13
$\hat{\beta}_{TSS}$	0.1317	0.4332	0.2917	0.07
$\bar{R}^2$	0.43			
SSE	0.0183			
$ln\ L$	-24.063			

Table 3.3 : Regression results for second set of samples

	Estimates	S.E.	t-ratio	Partial Correlation
$\hat{\beta}_0$	-6.9431	22.80	0.3045	-0.08
$\hat{\beta}_{TF}$	7.2749	3.053	2.381	0.44
$\hat{\beta}_{TR}$	0.7900	0.2546	3.103	0.61
$\hat{\beta}_B$	-0.3154	0.2380	-1.325	-0.31
$\hat{\beta}_{TLC}$	3.3711	1.029	3.277	0.63
$\hat{\beta}_{TSS}$	-0.1288	0.645	0.1997	-0.05
$\bar{R}^2$	0.49			
SSE	0.061			
$ln\ L$	-34.93			
$\hat{\beta}_0$	-14.209	12.84	1.106	-0.27
$\hat{\beta}_{TF}$	6.1134	2.675	2.285	0.50
$\hat{\beta}_{TR}$	0.1965	0.1393	1.317	0.31
$\hat{\beta}_B$	-0.1632	0.1390	1.174	-0.28
$\hat{\beta}_{TLC}$	0.6597	0.6911	0.9545	0.23
$\hat{\beta}_{TSS}$	0.4112	0.3565	1.153	0.28
$\bar{R}^2$	0.51			
SSE	0.016			
$ln\ L$	-20.60			

Table 3.4 : Regression result for third set of samples

	Estimates	S.E.	t-ratio	Partial Correlation
$\hat{\beta}_0$	3.976	8.238	0.4827	0.12
$\hat{\beta}_{TF}$	15.441	9.844	1.568	0.28
$\hat{\beta}_{TR}$	1.3243	0.5777	2.298	0.49
$\hat{\beta}_B$	-0.2173	0.3307	0.6571	-0.15
$\hat{\beta}_{TLC}$	1.1139	1.652	0.6744	0.16
$\hat{\beta}_{TSS}$	0.0863	0.1378	0.6261	0.15
$\bar{R}^2$	0.36			
SSE	0.0768			
$ln\ L$	-42.50			
$\hat{\beta}_0$	7.1877	4.552	1.579	0.35
$\hat{\beta}_{TF}$	5.1622	2.357	2.1735	0.56
$\hat{\beta}_{TR}$	0.4288	0.3196	1.341	0.30
$\hat{\beta}_B$	-0.1589	0.1892	0.8401	0.19
$\hat{\beta}_{TLC}$	0.5651	0.8936	0.6324	0.15
$\hat{\beta}_{TSS}$	0.0285	0.0759	0.3753	0.09
$\bar{R}^2$	0.39			
SSE	0.0205			
$ln\ L$	-26.84			

Table 3.5 : Regression result for fourth set of sample

	Estimates	S.E.	t-ratio	Partial Correlation
$\hat{\beta}_0$	-53.804	23.63	2.277	-0.52
$\hat{\beta}_{TF}$	2.5143	4.858	0.5175	0.10
$\hat{\beta}_{TR}$	-0.1377	0.5100	0.2700	0.07
$\hat{\beta}_B$	-1.0420	0.5039	2.068	0.48
$\hat{\beta}_{TLC}$	-2.3556	1.958	1.203	0.31
$\hat{\beta}_{TSS}$	1.2230	0.5470	2.236 0.51	
$\bar{R}^2$	0.51			
SSE	0.0504			
$ln\ L$	-28.7228			
$\hat{\beta}_0$	-16.467	12.30	1.338	0.34
$\hat{\beta}_{TF}$	1.3830	3.586	0.3857	0.10
$\hat{\beta}_{TR}$	-0.0897	0.3583	0.4456	-0.09
$\hat{\beta}_B$	0.6232	0.2685	2.321	0.53
$\hat{\beta}_{TLC}$	-0.9274	1.025	0.9046	-0.24
$\hat{\beta}_{TSS}$	0.4170	0.2815	1.481	0.37
$\bar{R}^2$	0.49			
SSE	0.011			
$ln\ L$	-14.38			

Table 3.6 : OLS and Ridge estimates for Data Set 3.

Estimates	OLS	Ridge
$\hat{\beta}_0$	2.4646	2.34
	(1.906)	(30.00)
$\hat{\beta}_{TF}$	-0.135	0.179
	(1.749)	(4.74)
$\hat{\beta}_{TR}$	0.121	0.13
	(1.430)	(5.43)
$\hat{\beta}_{TLC}$	0.476	0.475
	(1.54)	(7.56)
$\hat{\beta}_B$	-0.956	-0.955
	(1.767)	(2.767)
$\bar{R}^2$	0.59	0.74
SSE	1.307	0.1794
$ln\ L$	-40.76	-8.784
AIC	-2.614	-3.67
FPE	-0.569	-0.269
$\hat{\beta}_{dummy}$	0.1925	0.965
	(8.602)	(29.60)

Table 3.7 : Prediction Result

Observed		Ridge Prediction		LAD Prediction	
First set	Last set	First set	Last set	First set	Last set
3	3	2.945	3.012	2.935	3.226
3	3	2.976	3.002	2.937	2.957
3	2.6	3.025	2.70	3.04	2.634
2.6	2.6	2.531	2.641	2.473	2.712
3	2.6	3.021	2.680	3.20	3.004
3.4	3	3.31	3.002	3.082	2.740
2.6	3	2.496	3.012	2.438	2.941
3.4	3	3.297	2.977	3.118	2.860
3	2.6	2.989	2.611	2.99	2.610
3	3	3.011	3.012	3.021	2.923
Sum of Sqr. of Prediction Error		0.0392	0.0192	0.2733	0.3266

Table 3.8 : OLS results for individual equations

Equation for Q	Estimates	t-ratio
β (intercept)	7.1895	1.59
$\beta(TF)$	4.35	1.60
$\beta(TR)$	0.4044	2.90
$\beta(BR)$	-0.1599	1.61
$\beta(TSS)$	-0.1768	1.54
R^2	0.16	
SSE	28.48	
ln L	-174.614	
Equation for V		
β'(intercept)	5.9911	4.44
$\beta'(BR)$	-0.0077	0.449
$\beta'(TLC)$	-0.035	0.2788
$\beta'(S)$	0.3751	23.58
R^2	0.58	
SSE	18.85	
ln L	-172.614	

Table 3.9 : EGLS Estimates of regression coefficients and the values of test statistics

Coefficients/Test Values	Estimates	t-ratio
$\beta(TF)$	5.4966	1.965
$\beta(TR)$	0.4884	3.075
$\beta(BR)$	-0.1896	1.962
$\beta(TSS)$	-0.1963	1.470
$\beta'(BR)$	-0.0043	0.2678
$\beta'(TLC)$	-0.0804	0.5991
$\beta'(S)$	0.4081	24.659
R^2	0.78	
χ^2	139.05 (7 d.f.)	
LM test	15.827 (1 d.f.)	
LR test	29.552 (1 d.f.)	
ln L	-170.61	

CHAPTER - 4

SMALL SAMPLE ESTIMATION PROBLEM WITH BOX-COX TRANSFORMATION : APPLICATION TO TEA QUALITY ASSESSMENT DATA

4.1 Introduction

In the last chapter we have used only linear models to associate the biochemical information with Tea Taster's choices made on tea samples. In statistical studies linear models are applied with the assumption of homogeneity of variance, simplicity of structure for the expected value of the response variable and approximate normality of the additive errors. The independence of errors is also assumed.

But it may not always be possible to satisfy the above mentioned requirements in the original scale of measurement of the response variable. There may be inherent non-linearity in the data. In such a situation a non-linear transformation of the response variable may yield homogeneity of variance and, at least approximately, normality of the error. A better fit may thus be obtained.

In this chapter we explicitly deal with the possible non-linearity in a tea quality assessment data. The Box-Cox transformation technique is applied to a particular data set to achieve better fit. A vast literature is available on the transformation techniques (Atkinson, 1985), especially on Box-Cox technique (Box et al., 1964). Section 4.2 presents a brief discussion on the needs for transformation from application point of view. A brief review of literature on Box-Cox transformation and on related studies is presented in section 4.3. There are two approaches of estimation in the Box-Cox transformation model – (i)

maximization of likelihood function for the transformed model, (ii) minimization of SSE for the normalized transformation model. We note here that the inference procedures in the Box-Cox transformed model are basically based on large or moderately large samples (Carrol and Ruppert 1987, Atkinson 1985). Since in our quality assessment study, we deal with small sample size, it is important to investigate the performance of the usual estimation approaches with small sample data. An alternative non-linear least square method of estimation for Box-Cox transformation model is proposed, which is expected to give more robust estimates of parameters compared to those obtained through the usual estimation procedures. However, the inference procedure and other statistical properties of the proposed method of estimation are not discussed. In section 4.4 we try to study different possible estimation problems in the transformation model with small samples. The usual inference procedures for the transformation model are briefly discussed in section 4.5. The analysis of the Data Set 4 using different estimation procedures is presented in section 4.6.

4.2 Motivation Behind Nonlinear Transformation

As mentioned in the introductory section, if the fundamental assumptions behind a linear regression model is not satisfied, a nonlinear transformation of the response variable may reasonably meet the homogeneity and approximate normality requirements. Given a data set, the usefulness of a transformation may be indicated by empirical evidences. The non-negativity of the response variable may be one indication (Atkinson 1985). In this case the log of response is likely to be more close to normality. However, if the values of response are far from zero and variation among the values of response is relatively small, the transformation may have little effect.

To see whether there is any outliers one may use a normal plot of the residuals before and after transformations. The presence of outliers or, more particularly, the departures from the assumptions in the residuals, is sometimes an indication of the need for transformation. But, we note here that, if the presence of two or three outliners is an indication

of the need for a transformation, we may delete the outliers from the data set and then run the linear regression. Deletion of the outliers may be possible in case of large samples and if the deleted observations do not effect much in the analysis. But if the inclusion of all the observations is very much necessary for the analysis, and particularly, if the sample size is small, then we can not afford to delete the outlying observations. In such a situation, we have no way out but to opt for a suitable transformation.

Of many transformation possibilities, the most popular one is the parametric family of transformations analyzed by Box and Cox (1964), which brings out the choice of a transformation within the framework of standard statistical theory. In the following section we briefly present different aspects of Box-Cox transformation.

4.3 The Box-Cox Transformed Linear Model : A Brief Review

Box-Cox (1964) proposed the following parametric family of power transformation

$$y(\lambda) = \begin{cases} \frac{y^{\lambda}-1}{\lambda} & (\lambda \neq 0) \\ log\ y & (\lambda = 0) \end{cases} \tag{4.1}$$

In the absence of a transformation we have $\lambda = 1$ and the value of the transformation for $\lambda = 0$ is found as the limit of (4.1) as $\lambda \to 0$. Thus the function is continuous at '0' as $lim_{\lambda \to 0} \frac{y^{\lambda}-1}{\lambda} = ln\ \lambda$.

A Box-Cox transformed linear model is a model that usually takes the following form: for some unknown real value λ and $i = 1, 2, \ldots, n$,

$$y_i(\lambda) = \beta_1 + \beta_2\ x_{i2} + \ldots + \beta_k\ x_{ik} + e_i\ , \tag{4.2}$$

where e_i's are $iid\ N(0, \sigma^2)$. The intention is that, for some $\lambda, E\ [y(\lambda)] = X\beta$ with $y(\lambda)$ satisfying the condition of variance homogeneity, independence and additivity. Also it is expected that, to an adequate degree of approximation, $y(\lambda)$ will be normally distributed.

The power transformation (4.1) is only one of many parametric families of transformation developed by different researchers. A detailed account of such transformations is presented in Atkinson's book (1985). There is an extensive literature on the estimation

aspects as well as the problems in application of Box-Cox transformation (for instance, Egy and Lahiri 1979; Savin and White 1978). Collins (1991) reviews several techniques useful for forming point and interval prediction in regression models with Box-Cox transformed variables. Monte Carlo studies are made to examine the small sample accuracy using the techniques proposed. A related study is due to Carrol and Ruppert (1981), who study the cost, in terms of forecast mean squared error, of estimating λ.

Carrol (1982) discusses how one can test the regression parameters within the context of the Box-Cox power transformation family. He proposes a simple conditional test which consists of estimating the correct scale and than proposes the use of usual linear model F-test in this estimated scale. He investigates situations in which this test has the correct level asymptotically as well as comparable power to Wald's test or the LR test. Box and Cox (1982) and Hinkley and Runger (1984) take a conditional approach that essentially says that one should make conditional inference for an appropriately defined regression parameter, conditioned on the data based $\hat{\lambda}$. Cox and Reid (1987) arrive at a similar conclusion to that of Hinkley et al.(1984). Recently Chen and Lockhart (1997) study the Fisher information matrix and in particular, it's inverse, for unknown parameters in the likelihood based analysis of Box-Cox model. They discuss the inference problem associated with β when λ is estimated from the data. Both conditional and unconditional inference procedures are studied.

We note here that most of the studies based on Box-Cox model, in our knowledge, deal with the large sample problems. We do not have specific information about any study dealing with the likely problems is applying Box-Cox transformed linear model to small sample data. Certain complexities may arise in making inference about β when λ is estimated from the data, in case of small sample size. In the following section we try to make some comments on the possible problems that may arise in this situation.

4.4 Some Comments On The Estimation Problem

The likelihood function (LF) for the transformation model (4.2) may be written as

$$L = \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n exp\left[-\frac{1}{2\sigma^2}(y(\lambda) - X\beta)'(y(\lambda) - X\beta)\right] J \ , \tag{4.3}$$

where $J = (\Pi y_i)^{\lambda-1}$ is the term due to Jacobian of transformation. The ML estimate of β for given λ, which we denote by $\hat{\beta}(\lambda)$, is the least squares estimate given by $\hat{\beta}(\lambda) = (X'X)^{-1}X'y(\lambda)$. The sum of squares of errors of the $y(\lambda)$ is

$$\begin{aligned} SSE(\lambda) &= y(\lambda)'\{I - X(X'X)^{-1}X'\}\, y(\lambda) \\ &= y(\lambda)'(I - H)\, y(\lambda) = y(\lambda)'\, M\, y(\lambda). \end{aligned} \tag{4.4}$$

The ML estimate of the residual variance is $\hat{\sigma}^2(\lambda) = SSE(\lambda)/n$. For fixed λ, the profile loglikelihood (using the estimators of β and σ^2), apart from the constant, may be written as

$$l_{max}(\lambda) = -\frac{n}{2}\, ln\, \hat{\sigma}^2(\lambda) + ln\, J,$$

which is a function of λ, and clearly depends on $SSE(\lambda)$ and the Jacobian J.

An equivalent form for $l_{max}(\lambda)$ may be found by using the normalized transformation

$$Z(\lambda) = y(\lambda)/J^{\frac{1}{n}}.$$

For the power transformation (4.1), we have $\frac{\partial y_i(\lambda)}{\partial y_i} = y_i^{\lambda-1}$, so that $log\ J = (\lambda - 1)\ \sum_i log\ y_i$. Under this transformation we see that the y_i values are divided by their geometric mean (GM). Let the GM of y_i values be $G(y)$. Then the normalized power transformation is

$$\begin{aligned} Z(\lambda) &= \frac{y^{\lambda}-1}{\lambda[G(y)]^{\lambda-1}} \qquad (\lambda \neq 0) \\ &= G(y)\ log\ y \qquad (\lambda = 0). \end{aligned} \tag{4.5}$$

Apart from a constant, the partially maximized log-likelihood of the observations can be written as

$$l_{max}(\lambda) = -\frac{n}{2}\, ln\, [R(\lambda)/n]$$

where $R(\lambda) = Z'(\lambda)\ M\ Z(\lambda)$ is the SSE of the $Z(\lambda)$.

Here maximization of the likelihood function becomes equivalent to minimization of RSS of the transformed values. This is true as long as GM of the response values are regarded as constant. In fact a simple modification of the transformed estimates would give the estimates for original parameters. The loglikelihood function in original parameters, without the constant term is

$$-\frac{n}{2}\ ln\ \sigma^2 - \frac{1}{2\sigma^2}\,[y(\lambda) - X\beta]'\,[y(\lambda) - X\beta] + (\lambda - 1)\ \sum\ ln\ y_i\ .$$

Now suppose each response value is multiplied by a constant C and we maximize the loglikelihood function to estimate the parameters. In this case we have

$$\begin{aligned}
&-\tfrac{n}{2}\ ln\ (\sigma')^2 - \tfrac{1}{2(\sigma')^2}\ \Sigma\left(\tfrac{C^\lambda y_i^\lambda - 1}{\lambda} - \beta_0' - \beta_1'\ x_{1i} - \ldots - \beta_k'\ x_{ki}\right)^2 \\
&+(\lambda - 1)\ \Sigma\ ln\ (Cy_i) + n\ ln\ C \\
&= -\tfrac{n}{2}\ ln\ (\sigma')^2 - \tfrac{1}{2(\sigma')^2}\ \Sigma\left(\tfrac{C^\lambda y_i^{\lambda-1} - C^\lambda + C^\lambda - 1}{\lambda} - \beta_0' - \ldots - \beta_k'\ x_{ki}\right)^2 \\
&+(\lambda - 1)\ \Sigma\ ln\ (Cy_i) + n\ ln\ C \\
&= -\tfrac{n}{2}\ ln\ (\sigma')^2 - \tfrac{1}{2(\sigma')^2}\ \Sigma\left[C^\lambda\ y_i^{(\lambda)} + C^\lambda - \beta_0' - \ldots - \beta_k'\ x_{ki}\right]^2 \\
&+(\lambda - 1)\ \Sigma\ ln\ (Cy_i) + n\ ln\ C \\
&= -\tfrac{n}{2}\ ln\ (\sigma')^2 - \tfrac{1}{2(\sigma')^2}\ \Sigma(C^\lambda)^2\left[y_i^\lambda - \tfrac{\beta_0' - C^{(\lambda)}}{C^\lambda} - \tfrac{\beta_1'}{C^\lambda}\ x_{1i} - \ldots - \tfrac{\beta_k'}{C^\lambda}\ x_{ki}\right]^2 \\
&+(\lambda - 1)\ \Sigma\ ln\ (Cy_i) + n\ ln\ C \\
&= -\tfrac{n}{2}\ ln\ (\sigma')^2 - \tfrac{(C^\lambda)^2}{2(\sigma')^2}\ \Sigma\left[y_i^\lambda - \beta_0 - \beta_1\ x_{1i} - \ldots - \beta_k\ x_{ki}\right]^2 \\
&+(\lambda - 1)\ \Sigma\ ln\ (Cy_i) + n\ ln\ C \\
&= -\tfrac{n}{2}\ ln\ (\sigma)^2 - \tfrac{1}{2(\sigma)^2}\ \Sigma\left[y_i^{(\lambda)} - \beta_0 - \beta_1\ x_{1i} - \ldots - \beta_k\ x_{ki}\right]^2 \\
&+(\lambda - 1)\ \Sigma\ ln\ (y_i)\ ,
\end{aligned}$$

$$\text{where } (\sigma)^2 = (\sigma')^2/(C^\lambda)^2,\ \ \beta_0 = \frac{\beta_0' - C^{(\lambda)}}{C^\lambda},\ \ \beta_1 = \frac{\beta_1'}{C^\lambda}, \ldots \beta_k = \frac{\beta_k'}{C^\lambda}. \qquad (4.6)$$

Thus the original estimates can be obtained from (4.6). The estimate of λ does not differ. The same comment is applicable when $C = \frac{1}{GM}$ and GM is regarded as constant. But GM being a function of all the y_i values can not always be regarded as constant. Let

us now try to develop the form of likelihood function under varying GM. The underlying transformation is

$$Z_i = y_i \left(\Pi\, y_l\right)^{\frac{1}{n}} \tag{4.7}$$

The derivatives of the Jacobian of transformation may easily be obtained as

$$\frac{\partial Z_i}{\partial y_j} = \begin{cases} -\frac{y_i}{n y_j}\ (\Pi y_l)^{-\frac{1}{n}}, & i \neq j \\ -\frac{(\Pi\, y_l)^{-\frac{1}{n}}}{n} + (\Pi\, y_l)^{-\frac{1}{n}}, & i = j \end{cases} \tag{4.8}$$

It may easily be examined that the determinant of this matrix is zero :

$$\left|\frac{\partial(Z_1, Z_2 \ldots Z_n)}{\partial(y_1, y_2 \ldots y_n)}\right| = \frac{(\Pi y_i)^{-1}}{n^n} \begin{vmatrix} 1-n & \frac{y_1}{y_2} & \ldots & \frac{y_1}{y_n} \\ \frac{y_2}{y_1} & 1-n & \ldots & \frac{y_2}{y_n} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \frac{y_n}{y_1} & \frac{y_n}{y_2} & \ldots & 1-n \end{vmatrix}$$

$$= \frac{(\Pi y_i)^{-2}}{n^n} \begin{vmatrix} y_1(1-n) & y_1 & \ldots & y_1 \\ y_2 & y_2(1-n) & \ldots & y_2 \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ y_n & y_n & \ldots & y_n(1-n) \end{vmatrix}$$

$$= \frac{(\Pi y_i)^{-2}}{n^n} \begin{vmatrix} (1-n) & 1 & \ldots & 1 \\ 1 & 1-n & \ldots & 1 \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ 1 & 1 & \ldots & 1-n \end{vmatrix} = 0\ ,$$

since the sum of the rows in the matrix is equal to zero vector.

Instead of transformation twice (first by λ-transformation and than dividing y_i values by their GM), we may consider the normalized power transformation (4.5). This also

suffers from the same defect, as expected. The Jacobian of transformation is again zero. The result has a serious implication on the estimate of SE of the parameters. In this case the estimates obtained by minimizing ESS of the transformed response values need not be same as those obtained by maximizing the loglikelihood function. But it is always possible to obtain one set of estimates from the other. So these are equivalent procedures as long as we are interested only is estimating the parameters. However, the SE's of the estimates will differ for these two procedures, as because the GM should not be regarded as a constant. The problem would be more serious in case of small samples, and further investigation is necessary to study the likely complexities in estimating the SE and to go for inference. Asymptotically, however, both the procedures will result in the same value of S.E.

When the minimum of y values (y_{min}, say) approaches zero, the likelihood function defined in (4.3) or the corresponding loglikelihood, when maximized, becomes unbounded. The minimum value of ESS of the transformed observations becomes zero. This is because GM is very much sensitive to small observations. As y_{min} approaches zero, the GM also approaches zero. Change in GM value is significant even for a small change in y_{min}, when y_{min} is close to zero. Thus GM can not be regarded as constant.

The Box-Cox power transformation discussed above is based on the implicit assumption that the apparent origin of the response variable is a true lower limit. In general linear model, the subtraction of a constant may be required before taking transformation. In this case only the value of the intercept is affected leaving other aspects of the model unchanged. But the effect of changing the values of the response variable by shifting the origin, before transformation, may be serious. This may alter many aspects of the fitted model, including the estimate of λ. The empirical necessity for such shift in origin in the transformation model is discussed by Atkinson (1985, p. 184). The normalized form of the transformation introducing a shift parameter μ, compared to (4.5) is

$$Z(\lambda,\mu) = \begin{cases} \frac{(y+\mu)^{\lambda}-1}{\lambda[G(Y+\mu)]^{\lambda-1}} & (\lambda \neq 0) \\ \\ ln\ (y+\mu)\ G(y+\mu) & (\lambda = 0), \end{cases} \tag{4.9}$$

where $G(y+\mu)$ is the geometric mean of the constructed response variable after shift is origin.

In this formulation (4.9) also the problem of estimation become severe. If μ approaches - y_{min}, there is at least one value of $y+\mu$ which becomes very small, so that $G(y+\mu)$ also becomes very small. In this situation the likelihood function becomes unbounded or the ESS zero.

One way to resolve the problem of shifted origin, we think, is to obtain a nonlinear least square solution by minimizing $\sum(y_i - \hat{y}_i)^2$, where $\hat{y}_i$ is the estimate of the untransformed y_i. For the transformation with shift parameter :

$$y(\lambda,\mu) = \begin{cases} \frac{(y+\mu)^{\lambda}-1}{\lambda} & (\lambda \neq 0) \\ \\ ln\ (y+\mu) & (\lambda = 0), \end{cases}$$

we may obtain $\hat{y}_i$ as

$$\hat{y}_i = \begin{cases} [\lambda(x_i'\beta+1]^{\frac{1}{n}} - \mu & (\lambda \neq 0) \\ \\ Exp\ (x_i'\beta) - \mu, & (\lambda = 0). \end{cases}$$

The above formulation is straightforward from the relationship $\frac{(y_i+\mu)^\lambda-1}{\lambda} = x_i'\beta+e_i$ *for* $\lambda \neq 0$.

4.5 The Testing Aspects in Transformation Model

Extensive studies have been made by researchers to investigate the testing aspects of regression parameters within the context of Box-Cox transformation family. To refer a few among many others : Carrol (1982), Box and Cox (1982), Hinkley and Runger (1984), Cox and Reid (1987). Recently Chen and Lockhart (1997) discuss the inference problem associated with β when λ is estimated from the data.

The main aim is to test whither a transformed model fit better in comparison to the linear model for a given data set. Thus the null hypothesis is H_0 $\lambda_0 = 1$ against the alternative H_1 $\lambda_0 \neq 1$. The likelihood based test is used in general. To compare the likelihood for various values of λ, $L_{max}(\lambda)$ can be plotted over a range of plausible values. An approximate $100(1-\alpha)$ per cent confidence region for λ is found for those values for which

$$2\left[L_{max}(\hat{\lambda}) - L_{max}(\lambda)\right] \leq \chi^2_{1,\alpha} \tag{4.10}$$

We note here that the SSE, $S(\lambda) = y(\lambda)'[I - X(X'X)^{-1}\ X']\ y(\lambda) = y(\lambda)'\ My(\lambda)$, of the transformation (4.1), can not be used to compare the adequacy of the models for various values of λ. This is because of the change of scale on transformation. However, $R(\lambda) = Z(\lambda)'\ MZ(\lambda)$, the SSE based on the normalized transformation (4.9), can be used as the basis of an approximate method for making such comparisons. But in case of small samples, this approach also has a serious drawback. This is because the sampling variability of the geometric mean of y, which is included in $Z(\lambda)$, is ignored. In case of small samples we can not consider $GM(y)$ to be constant. The problem becomes severe if the minimum values of y approaches zero. However, asymptotically sampling fluctuation of $GM(y)$ becomes less important and the distribution of $Z(\lambda)$ converges to that of $y(\lambda)$.

That is why, in small sample situation, we think that the comparison of the confidence regions for λ based on $S(\lambda)$ and $R(\lambda)$ should be compared. In terms of $R(\lambda)$, the confidence

region for λ includes all those values for which

$$R(\lambda) - R(\hat{\lambda}) \leq (\chi^2_{1,\alpha}/n)\, R(\hat{\lambda}) \tag{4.11}$$

As noted by Atkinson (1985), the confidence regions based on $S(\lambda)$ and $R(\lambda)$ are asymptotically equivalent. In case of small sample, the confidence interval (4.11) will be broader due to term 'n' is the denominator, as compared to the region (4.10).

4.6 Analysis of Data Set 4 Using Box-Cox Transformation Model

Here our primary aim is to associate the biochemical information for 23 CTC tea samples with the sensory evaluation made by a single Taster. We first fit the linear regression model. The estimates of regression parameters along with other information about regression diagnostics are presented in Table 4.1. As evident from the table, the OLS offers a very poor fit with low adjusted R^2 (= 0.28) and hence with comparatively higher value of SSE. All the biochemical quality parameters except CF are statistically insignificant, as evident from the t-ratios (or p-values). To test the normality of residuals we apply the χ^2 test of normality of residuals. The hypothesis of symmetric residuals is rejected at 5% level of significance.

Guided by asymmetry of residuals and the poor OLS fit, we try the Box-Cox transformed linear model. We first obtain the estimates of parameters by maximizing the likelihood function (4.3) of the transformed model. The estimates along with information required for diagnostics are presented in Table 4.2. Apparently this gives a better fit in the sense of higher value of $\bar{R}^2$ (= 0.51) and the value of SSE (10.2872). Only the quality parameter CF is statistically significant along with the intercept term, as evident from the conditional t-ratios. This conditional t-ratio is referred in section 4.3. Comparing the loglikelihood values of Table 4.1 and 4.2, we see that the twice of the difference between the loglikelihood values is greater than 3.83, the 5% value of χ^2 with 1 df. Thus using the confidence region (4.10), we may reject the null hypothesis $H_0 : \lambda = 1$ at 5% level.

Apart from SSE and loglikelihood, we may also use some information criteria to compare the fit. Some popular likelihood based information criteria are : Akaike Final Prediction Error (FPE), Akaike Information Criteria (AIC) and Shibata Criteria (SC). A detailed discussion on these information criteria is available in the book of Judge et al. (1985). Lower the values of these criteria better is the fit. In our case, the values of these information criteria are much less for transformation model as compared to those for the general liner model.

The values of loglikelihood and SSE $[S(\lambda)]$ for various values of λ, around the optimum value (λ = -1.79), are presented in Table 4.3. Note that because of change of scale on transformation, the SSE $[S(\lambda)]$ values can not be used to compare adequacy of the model for various values of λ. To compare the various values of λ, $L(\lambda)$ can be plotted over a range of plausible values. An approximate confidence region for λ, using (4.10), in this case is (-1.75, -1).

Another approach in estimation is to minimize SSE for the normalized transformation $Z(\lambda)$ defined on (4.5). Here $R(\lambda) = Z(\lambda)'(I - X'(X'X)^{-1}X]\ Z(\lambda)$ is the SSE. The $R(\lambda)$ values for various values of λ may be considered as the basis of an approximate method of making comparison for model adequacy. But as we have mentioned in section 4.4, the sampling variability of $GM(y)$ may seriously affect the estimation in case of small samples. In this situation, maximization of loglikelihood and minimization of SSE $[R(\lambda)]$ will not give similar results, λ being estimated from the data. For small sample, the estimates obtained using both the approaches, which are otherwise equivalent, should not be compared. However, asymptotically sampling fluctuations in $GM(y)$ is not important and the distribution of $Z(\lambda)$ converges to that of $y(\lambda)$.

The values of SSE for various values of λ, obtained through minimization of $R(\lambda)$, are presented in Table 4.4. The optimum value of $\hat{\lambda}$ is -1.7894 with SSE = 9.0266. It is interesting to observe that although the estimated values of λ are approximately equal for both the methods of minimum SSE and maximum loglikelihood, the SSE values differ. The estimated β coefficient, obtained by normalized transformation, are: β_0 = 133.3119, β_{TF} = 0.6517, β_{TR} = 0.1242, β_{CAF} = 0.0897, β_{CF} = -0.6298, β_A = -0.5943. These $\hat{\beta}$ values

are different from those presented in Table 4.2. One of the reasons for these differences in the estimates obtained is the effect of the sampling variability of $GM(y)$. The plot of SSE against λ is presented in Figure 4.1 with 95 per cent confidence interval line. The estimated values of χ^2 (using $R(\lambda)$) is 4.4046, rejecting the null hypothesis of $\lambda = 1$ at 5% level of significance. The confidence region for λ takes all those values of λ for which the SSE values are below the confidence like in Figure 4.1. An approximate 95 per cent confidence interval for λ is (-1.79, 0.9474), which is much wider than the confidence interval obtained through maximization of loglikelihood.

Thus we observe that in case of small samples, the usual estimation approaches in Box-Cox transformed model may give misleading results. The problem may be more severe if the minimum value of y approaches zero after transformation, when we use the normalized version $Z(\lambda)$. In this situation the non-linear least squares (NLLS) approach, proposed in section 4.3, is expected to give more stable estimates. This approach is likely to be more robust as compared to estimation from $Z(\lambda)$. With this approach, the estimated value of λ is 0.6703 with SSE = 9.0233. The percentage of variation explained is 57.1. A fall in the value of SSE as compared to those in Table 4.1 and 4.2 may be observed. Note that $\hat{\lambda}$ = 0.6703 is within the confidence limit obtained through normalized transformation. The estimated values of regression coefficients, obtained using NLLS method are β_0 = 8.2586, β_{TF} = 0.1956, β_{TR} = 0.0493, β_{CAF} = 0.1516, β_{CF} = -0.3530 and β_A = -0.3421.

4.7 Concluding Remarks

On the basis of our study, we may say that the likelihood based approach of estimation in Box-Cox transformation model should not be adopted in case of small sample data. The approach of minimization of SSE for the normalized transformation is sensitive to sampling variability of $GM(y)$. The confidence region for λ is unduly broadened due to the effect of small sample size. The proposed non-linear estimation approach, which is very trivial, seems to be more robust than the usual estimation approaches. Moreover, since the problem of normalization of the transformed model is not involved in the estimation

process, it can be applied to small sample.

In this chapter we have not addressed the problem of subjectivity (error) associated with the response. As discussed in the last chapter, the response measurement error dose not pose much of statistical problems, apart from affecting the standard errors of the regression coefficients. However, we may theoretically formulate a transformation model incorporating the error component associated with the response variable. Such a formulation is presented as *Note1* in the concluding chapter.

Table 4.1 : OLS estimates and information for regression diagnostics for Data Set 4

	Estimates	S.E.	t-ratio	p-value
β_0	15.2959	6.3571	2.4061	0.027
β_{TF}	0.084	0.2388	0.3528	0.73
β_{TR}	0.1028	0.2684	0.3831	0.70
β_{CAF}	0.0938	0.2569	0.3651	0.71
β_{CF}	-0.5405	0.2101	-2.5723	0.02
β_A	-0.1962	0.1965	-1.0015	0.33
SSE	11.9310			
σ^2	0.664			
$\bar{R}^2$	0.28			
FPE	0.8108			
AIC	0.7233			
SC	0.8009			
χ^2 normal	11.6944			
ln L	-24.0818			

Table 4.2 : Estimates using ML method for transformation $y(\lambda)$

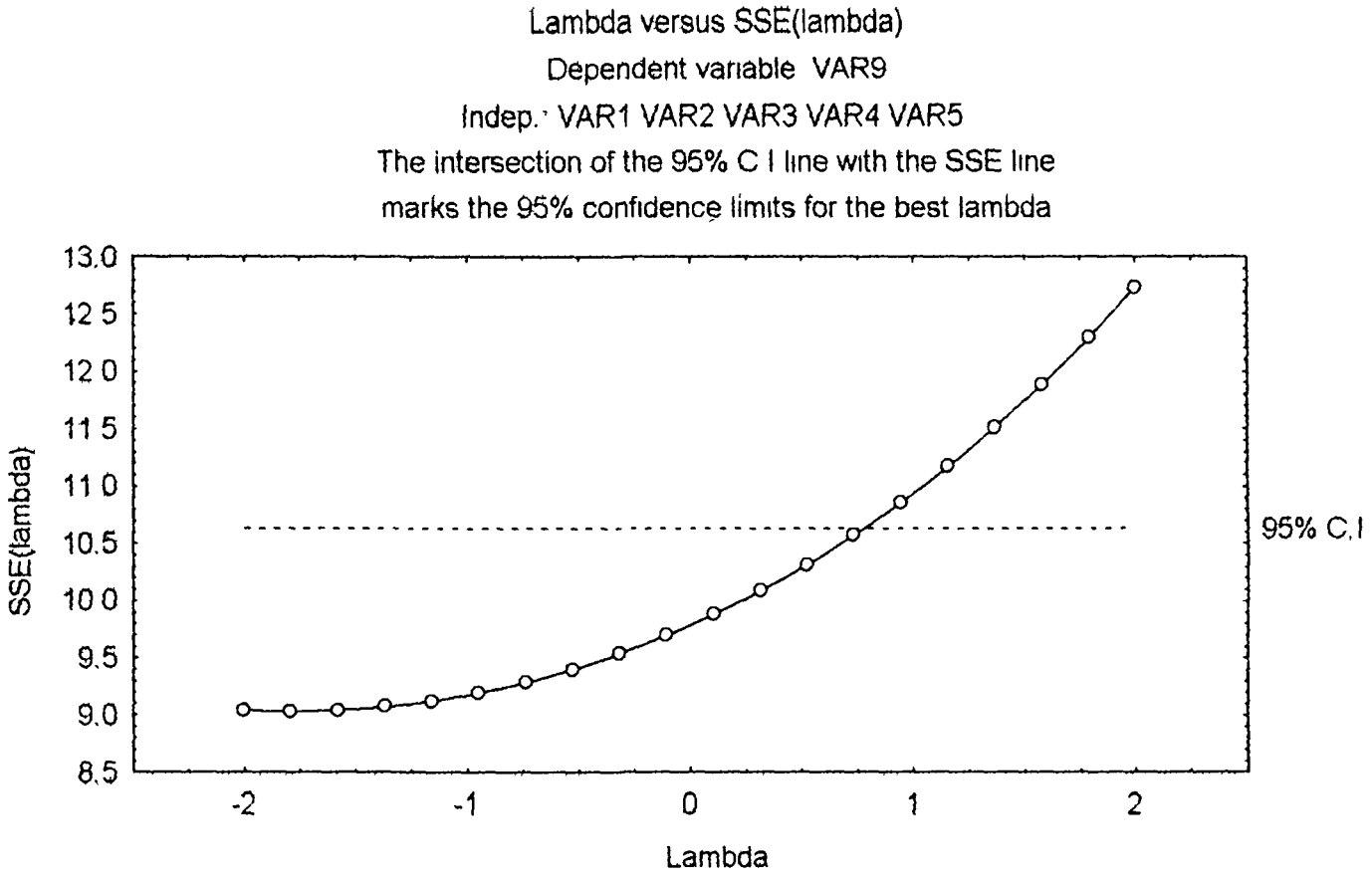
	Estimates	S.E.	Conditional t-ratio	p-value
β_0	0.5721	0.2493	22.95	0.00
β_{TF}	0.0281	0.0368	0.7631	0.45
β_{TR}	0.0054	0.0092	0.57%	0.57
β_{CAF}	0.0039	0.0345	0.1127	0.91
β_{CF}	-0.0256	0.0104	2.613	0.018
β_A	-0.0256	0.0216	0.9846	0.34
$\bar{R}^2$	0.51			
SSE	10.2872			
σ^2	0.4907			
FPE	0.0001			
AIC	0.0001			
SC	0.0001			
χ^2 normal	5.5573			
λ	-1.79			
ln L	-21.8795			

Table 4.3 : Log likelihood values along with SSE for different values of λ.

λ	ln L	SSE
-2.00	-21.8920	0.0007
-1.75	-21.8800	0.0019
-1.50	-21.9049	0.0052
-1.25	-21.9663	0.0139
-1.00	-22.064	0.0373
-0.75	-22.1974	0.1001
-0.50	-22.3661	0.2697
-0.25	-22.5995	0.7284
0	-22.8069	0.1973
0.25	-23.0776	0.5361
0.50	-23.3809	1.4607
0.75	-23.7159	3.9906
1.00	-24.0818	10.932

Table 4.4 : Values of λ and SSE for normalized transformation $Z(\lambda)$.

λ	SSE
-2.00	9.0365
-1.7894	9.0266
-1.5789	9.0373
-1.1579	9.1202
-0.9474	9.1927
-0.5263	9.4009
-0.3158	9.5376
-0.1053	9.6366
0.1053	9.8789
0.5263	10.3158
0.7474	10.8565
1.1579	11.1688
1.5789	11.884
1.7894	12.2919
2.00	12.7343


Lambda versus SSE(lambda)
Dependent variable VAR9
Indep.: VAR1 VAR2 VAR3 VAR4 VAR5
The intersection of the 95% C I line with the SSE line
marks the 95% confidence limits for the best lambda
13.0
12.5
12.0
11.5
11.0
10.5
10.0
9.5
9.0
8.5
SSE(lambda)
-2
-1
0
1
2
Lambda
95% C.I

CHAPTER - 5

ANALYSIS OF TASTERS' SENSORY SCORES

5.1 Introduction

In this chapter we explicitly deal with the sensory evaluations made by the Tea Tasters in assessing the quality of tea samples. Sensory analysis, where the Tasters/assessors give scores on a structured or non-structured scale for several attributes, is usually called sensory profiling or descriptive sensory analysis (Amerine et al.,1965). Sensory panel data, where the assessors evaluate different products in terms of one or more attribute(s), are often blurred by extensive individual variations. These individual variations arise purely from the subjectivity inherent in the process of sensory analysis. The subjectivity in the sensory evaluation may be analyzed in several ways.

The sensory panel data are often analyzed by ANOVA technique based on the raw data or by multivariate technique like principal component analysis after averaging over the assessors' scores (Martens, 1985). Some statistical studies have been made to handle the individual difference among the assessors in sensory profiling. One remarkable study in this line is due to Brockhoff et al. (1994). They discuss the linear variance component models which take into account the scale differences among assessors as well as the reproducibility aspect. They also address the problem of measuring assessor precision and propose testing procedures for the significance of difference among the assessors' error variances. Naes (1990) discusses the statistical analysis and interpretation of data from sensory analysis. Techniques are discussed to handle the differences among assessors in using the 'scale'. Some studies in this line, specific to the Tea Tasters' sensory data, have been made by Pal et al. (1997) and Paul (1998, 2000). Pal et al. (1997) discusses the problem of estimating the mean profile for quality attributes specific to CTC tea samples using the Tasters' sensory panel data. The heteroscedastic variance component models and the ML estimation procedures are discussed. A similar study addressing the problem of Tasters' precision is due to Paul (1998). The problem of detecting assessors' reproducibility is addressed by Naes (1999) and Paul (2000), among others. Naes and Solheim

(1998) have addressed the problem of individual variations using graphically oriented tools.

We aim at addressing the 'subjectivity' inherent in the choices made by the Tasters through modeling individual difference among the assessors. If a single Tester evaluates a set of tea sample, it is not possible to study the error associated with his choices. But if the same sample is evaluated by panel of Testers, we may statistically study the bias associated with the sensory scores. Thus, in a restricted sense, we may consider this as a problem of repeated measurements. In our study we consider ordinal choices (given on a structured scale), made by a panel of Testers independently on a tea sample, as the repeated measurements or repeated observations.

In reality, the sensory panel data are more complex and there are various fixed and interaction effects with several other combinations, which needs to be taken care of while understanding the data clearly. Brockhoff et al. (1994) has discussed several complex aspects of the sensory panel data. Replicated sensory scores on different quality attributes for a particular sample may be given by a panel of assessors. In this case each assessor may give scores on a sample for different quality attributes on different scales. The interaction effect(s) mentioned above includes 'different use of scale'. Also, the individual variances, measured using replications on the same experimental sample, may very among the assessors. There are several other aspects also. Thus formulation of a unified model for sensory profile data is rather a complicated task.

In a typical sensory panel data the main emphasis should be given to the following two problems.

1. The within assessor variability. It means the detection of differences in reproducibility among assessors. Here the assessor with good reproducibility can be distinguished from those who are more unreliable.

2. Detection and interpretation of differences or variation among the assessors. This is the between assessor variability.

Following Naes (1990) and Brockhaff et al.' (1994), we may formulate a model for a sensory panel data. Consider the following generic situation. Suppose there are r

assessors in the panel evaluating K quality attributes for n samples. Also there may be several replicates (say, L) available for each sample and each Taster. Thus without loss of generality, the data may be decomposed for particular attribute k by the following additive models

$$y_{ijkl} = \mu_k + \alpha_{ik} + \beta_{jk} + \lambda_{ijk} + u_{ijkl} \tag{5.1}$$

Here μ_k represents the overall mean effect for k^{th} attribute. α_{ik} represents the variation between i^{th} assessor's average score for k^{th} attribute and the overall average for the same attribute. The effect β_{jk} describes how average for j^{th} sample deviates from the overall average for the k^{th} attribute. The interaction effect λ_{ijk} represents the difference among assessors in differentiating among the samples. In the model (5.1), the individual differences among assessors are present in both the main effect α_{ik} and interaction effect λ_{ijk}. u_{ijkl} is the error component representing variation due to replicates under the same experimental condition.

In most applications, the researchers have been interested in the analysis and interpretation of the sample effects β_{jk}. If we are to model the individual differences among assessors, then the interest lies on the main effect α_{ik} and interaction λ_{ijk}. Generalization of (5.1) may be done and ANOVA can be performed to test the significance of different effects.

We note here that the database for our study does not allow us to go deep into different complex aspects of sensory panel data. We do not have replicated observations on each sample. Also, is almost all the cases, measurements are available only on one quality attribute (e.g. overall quality). Given the minimum information on the basis of available data, we restrict to the analysis of over simplified sensory experiments.

The basic data format may be identified tersely as "n samples $\times r$ measurements". We consider only the balanced and complete data. By balanced data we mean that the r occasions of measurement are the same for all experimental samples. By complete data we mean that measurements are available at each point of observations for each experimental sample. We do not consider the incomplete or missing data problems in our study.

In the next section we briefly discuss the assumptions on Tea Tasters' effects in connection with the ANOVA models. In section 5.3 we address the statistical aspects in

measuring the *reproducibility* in sensory panel data. We study the Tasters' reproducibility under heterogeneity. This is a theoretical extension of Naes's (1998) work. In section 5.4 we discuss a naive statistical approach to estimate the Tasters' bias. Formulations are done under the assumptions of both correlated and independent error components. The testing aspect is also discussed.

The statistical approach to modeling the individual differences among Tasters using one-way VC model is presented in section 5.5. Under the assumption of unequal error variances associated with the Tasters, the ANOVA and ML estimators of variance components for two models are presented in section 5.6 and section 5.7.

The two-way mixed effects variance component model and the estimation procedures are discussed in section 5.8. Both ML and ANOVA estimators of variance components and fixed effects parameter are obtained under the mixed-effects formulation. The mixed effects formulation is useful when we legitimately assume that the sample specific effects are fixed and the Tasters' effects are random. Analysis of the Data Sets 5 and 6 are presented in section 5.9.

5.2 Random or Fixed Assessors' Effects

In most ANOVA applications it is assumed that all the assignable effects arising from a typical sensory panel data analyses are fixed parameters. In sensory analysis this implies that the sample specific and assessor specific effects are fixed. For samples, this 'fixed' assumption may be logical if we are dealing with specific products or different category products belonging to the same class. Here the inference would be based on the specific set of samples. However, if the samples truly represent a particular product, then one may legitimately assume the sample specific effects as random. But for the assessors (Tasters) this is a more questionable assumption. If the hypothesis of interest is $H_0 = \alpha_1 = \ldots = \alpha_r = 0$, under fixed effects formulation, then the possible consequence is that the results refer only to differences among average levels taken over the actual assessors. These averages can be quite different from the averages taken over the whole population. This is actually more interesting aspect to investigate.

On the other hand, if we consider the Tasters as random representatives from a pop-

ulation of trained persons, the Tasters' effects as well as the interaction effects should be considered as random effects in the model. One may raise question against this random Tasters' effects on the ground that the selection of Tasters in not a random process; the Tasters have gone through training process. In our opinion, the number of Tasters in a sensory panel is likely to be quite close to the population of all Tasters. At least in case of tea industry, this is very much true. There are only a handful number of Tasters working in the tea industry of India. That is why, we consider the effects due to Tasters as random.

5.3 Detecting Tasters' Reproducibility Under Heterogeneity

A measurement is said to be reproducible if, on repetition under similar condition, it gives the same results; that is to say, if the variation between the measurements are small and negligible. To assess a change or variation in the sensory evaluation on a particular tea sample by the Taster, the reproducibility of measurements and the relevant factors affecting the results should be known. We may legitimately consider that the variability in sensory scores on a particular sample (product) is due to the subjectivity associated with Taster's choice. In general scientific activities, the variability of results may be due to technical reasons band to the analytic method and to the equipment. Variability may also be due to measurement technique and the measurers, as well as the testing environment. Therefore, a simple description of variability is insufficient and a proper analytic model is needed that can quantify the different sources of random variation.

The data analytic problem is : what kind of statistical model is best suited for studies on reproducibility of Tasters' scores on tea samples. In this section we consider a simple statistical tool to detect the differences in reproducibility among the Tasters, that is, to detect the within Taster variability. This is possible only if the replicated scores for each sample given by each Taster is available.

The problem of detecting the differences in the reproducibility among assessors is addressed by Tormod Naes (1998) and residual variances (MSE) and F-values arising from an one-way ANOVA model (under homoscedasticity) are considered as statistical tools to differentiate among the assessors in their capacity to assess differences among

samples. Some graphical techniques are also discussed.

A general approach in this line, due to Diggle at al. (1995), to the analysis of repeated measures assumes a general linear model for the mean vector of the measurements and a particular parametric correlation model for the variance matrix that incorporates three qualitatively different sources of variation - (i) a random variation between samples, (ii) a positive correlation between measurements on the same sample over time (autocorrelation), and (iii) a random variation among responses within an individual. This formulation enables an estimation of the major components of variance and thereby an assessment of the reproducibility.

For particular attribute and assessor, the simple one-way ANOVA model to study the reproducibility of assessors may be written as

$$y_{il} = \mu + \alpha_i + u_{il} \tag{5.2}$$

$$i = 1, 2, ...n,\ l = 1, 2, ...L_i,$$

where α_i is the effect due to i^{th} sample and u_{ij} is the replicate error. Here we assume unbalanced data setup as the number of replicates need not be equal for each sample. We may assume $u_{il} \sim iid(0, \sigma_i^2)$, as in practice the unequal variances and sample sizes (L_i) appear to be the rule rather than the exception.

Though the aim is to discuss the MSE and F values only under the above formulation, we like to introduce a brief discussion on the F- tests, which are claimed to be robust under heterogeneity. We note here that the hypothesis of interest in the one-way formulation, in general, is $H_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_n$ against $H_1 : \alpha_i \neq \alpha_k$ for atleast one pair (i,k), $i \neq k$. Under H_0, with the assumption of equal variances, the likelihood ratio F- test is based on the F statistic

$$F = \frac{(n-1)^{-1} \sum_i L_i (y_{i0} - y_{00})^2}{(L-n)^{-1} \sum_i \sum_j (y_{ij} - y_{i0})^2}, \tag{5.3}$$

where $L = \sum L_i$, $y_{i0} = L_i^{-1} \sum_j y_{ij}$ *and* $y_{00} = L^{-1} \sum_i L_i y_{i0}$. Note that the term in the denominator of (5.3) is the MSE. Again, if $\sigma_i^2 \neq \sigma_k^2$ for atleast one pair (i, k), $i \neq k$, then the Welch's (1951) robustified version of F is

$$W = \frac{(n-1)^{-1} \sum_i a_i (y_{i0} - y_{00})^2}{1 + [2(n-2)/(n^2-1)] \sum_i (1 - a_i/a)^2/(L_i - 1)} \tag{5.4}$$

where $a_i = L_i/A_i^2$, $A_i^2 = \frac{1}{L_i-1}\sum_j (y_{ij} - y_{i0})^2$, $a = \sum_i a_i$ *and* $y_{00} = \frac{1}{a}\sum_i a_i y_{i0}$.

Assuming normality and unequal variances, the null distribution of W can be approximated by the F-distribution with $(n-1)$ and k degrees of freedom, where k is determined by

$$k^{-1} = [\frac{3}{n^2-1}]\sum_i \frac{(1-a_i/a)^2}{L_i - 1}.$$

The differences between F and W is that denominator of F is based on the pooled sample variance, whereas, in the denominator of W, the variances of the n samples are considered separately.

The one-way ANOVA F- tests are known to be α-robust in case of unequal error variances if the L_i are all equal. However, the F- test is very sensitive to the heterogeneity of variances for unequal sizes and to long tailed distribution. Some well known robustified versions of F- tests are Kruskal-Wallis test, Welch test (1951), among others. Krutchkoff (1988) discusses some common misconception about the F- tests and provides a simulation based solution to overcome drawback of the tests. The failure of the assumption of equal variances can have serious effect on the power of F- test. Krutchkoff (1988) provides an extensive study on the power performance of F- tests.

Turning back to our main aim of studying the MSE and F values under heteroscedasticity, we note that a more sensible approach to obtain the between-sample and within-sample (error) sum of squares would be to consider the standardized values. Before writing the standardized values, we obtain the ML estimators of the location parameters and the variance components as follows:

$$p(y_{il}) = \frac{1}{\sqrt{(2\pi)}\sigma_i} exp[-\frac{1}{2\sigma_i^2}(y_{il} - \mu - \alpha_i)^2].$$

The loglikelihood function may be written as

$$l = -\frac{L}{2}ln(2\pi) - \frac{1}{2}\sum_i L_i\, ln\sigma_i^2 - \frac{1}{2}\sum_i\sum_j \frac{1}{\sigma_i^2}\,(y_{il} - \mu - \alpha_i)^2.$$

Differentiating this l with respect to the parameters and equating to zero, the ML

estimators of the location parameter and the variance components may be obtained as

$$\mu = \frac{\sum_i w_i\, y_{i0}}{\sum_i w_i} \text{ and } \alpha_i = y_{i0} - \mu,$$

where $w_i = L_i/\sigma_i^2$ (i = 1, 2, ... n) are the weights. Here the means are appropriately weighted by the sample sizes to variance ratios. The ML estimator of variance components may be obtained as

$$\sigma_i^2 = \frac{1}{L} \sum_i \sum_l (y_{il} - \mu - \alpha_i)^2.$$

Note that the ML estimators are consistent and BLUE also. We now consider the standardized between-sample sum of squares as

$$SS\alpha = SS\alpha(\sigma_1^2, \sigma_2^2, ..., \sigma_n^2) = \sum_i^n L_i y_{i0}^2/\sigma_i^2 - (\sum_i L_i y_{i0}/\sigma_i^2)^2/(\sum_i L_i/\sigma_i^2), \tag{5.5}$$

and the standardized error sum of squares may be written as

$$SSe = \sum_i L_i S_i^2/\sigma_i^2, \tag{5.6}$$

where $S_i^2 = L_i^{-1}\ \sum_j\ (y_{il} - y_{i0})^2$ is the sample variance (MLE).

Using the ML estimators, the standardized sum of square may be written as

$$SS\alpha = \sum_i \hat{w}_i\ (y_{i0} - \hat{\mu} - \hat{\alpha}_i)^2,$$

which easily simplifies to obtain (5.5). We may incorporate the ML estimator of σ_i^2 in (5.6) and obtain the value of SSe. The estimates of mean sum of squares would be

$$MSe = (L - n)^{-1} SSe \text{ and } MS\alpha = (n - 1)^{-1} SS\alpha\ .$$

Now we can construct the F- test also. Note that $L_i\ S_i^2/\sigma_i^2 \sim \chi^2_{L_i - 1}$ and consequently, $SSe = \sum_i L_i\ S_i^2/\sigma_i^2$ has a χ^2 distribution with (L - n) degrees of freedom. Further, it can be easily shown that $SS\alpha$ defined by (5.5) has an independent Chi-square distribution with (n - 1) df when H_0 is true. Hence, as in standard one-way ANOVA with equal variances, we get under H_0,

$$F^* = \frac{SS\alpha(n-1)^{-1}}{SSe(L-n)^{-1}} \sim F_{n-1, L-n}. \tag{5.7}$$

Thus the SSe defined in (5.6) or the MSe and the F- ratio defined in (5.7) may be used to study the reproducibility of the assessors. It would be more logical to use these MSe and F value, as the effect of unequal variances has been taken care of while estimating the parameters. Also the estimation of sums of squares using the ML estimates of implicit location parameters and the variance components is sensible since the properties of ML estimators are well defined under very broad conditions.

5.4 A Naive Approach To Estimate Tasters' Bias

5.4.1 Independent Errors

Suppose three independent measurements (in our case, the quality scores given independently for the same response Y, say, the overall quality). Each observation is made up of two components - the true value of the variable (y) and an error in measurement.

$$Y_{i1} = y_{i1} + u_i, \quad Y_{i2} = y_{i2}v_i, \quad Y_{i3} = y_{i3} + w_i \, .$$

We may assume zero means and at least approximate normality for the measurement errors. If follows :

$$E(Y_1) = E(Y_2) = E(Y_3) = 0, E(Y_1 - Y_2) = E(u - v) = 0, E(Y_1 - Y_3) = 0 \text{ and } E(Y_2 - Y_3) = 0.$$

For small samples, we may use the paired t- test to study the significance of difference between the mean values of Y_1, Y_2 and Y_3. If the hypotheses of differences are rejected (for at least one case), we may proceed to estimate the independent measurement error variances as follows :

$$V(Y_1) = V(y) + V(u), V(Y_2) = V(y) + V(v), V(Y_3) = V(y) + V(w);$$

$$\left.\begin{aligned} V(Y_1 - Y_2) &= V(u) + V(v) \\ V(Y_2 - Y_3) &= V(v) + V(w) \\ V(Y_1 - Y_3) &= V(u) + V(w) \end{aligned}\right\} \quad , \quad \left.\begin{aligned} V(Y_1) - V(Y_2) &= V(u) - V(v) \\ V(Y_2) - V(Y_3) &= V(v) - V(w) \\ V(Y_1) - V(Y_3) &= V(u) + V(w) \end{aligned}\right\}$$

Using the above systems of equations we may obtain :

$$\begin{aligned} V(u) &= \frac{V(Y_1-Y_2)+V(Y_1)-V(Y_2)}{2} \quad or \quad \frac{V(Y_1-Y_3)+V(Y_1)-V(Y_3)}{2} \\ V(v) &= \frac{V(Y_2-Y_3)-V(Y_2)+V(Y_3)}{2} \quad or \quad \frac{V(Y_1-Y_2)-V(Y_1)+V(Y_2)}{2} \\ V(w) &= \frac{V(Y_2-Y_3)-V(Y_2)+V(Y_3)}{2} \quad or \quad \frac{V(Y_1-Y_3)-V(Y_1)+V(Y_3)}{2} . \end{aligned}$$

Clearly two estimates of the variances of each error component are obtained. we may test whether these two estimates differ significantly or not. Also, it may be easily verified that $V(u) \leq V(v) \leq V(w)$ implies $V(Y_1) \leq V(Y_2) \leq V(Y_3)$ We note here that under normality assumption, we may simultaneously estimate the error variances and also the variance of y through likelihood function.

5.4.2 Correlated Error Components

We now estimate the error variances under the assumption that the measurement error are correlated. Such formulation is logical if we assume that the Tasters' scores, though given independently, may be influenced by some common market related factors also. Under the assumption of correlated errors, we may formulate the problem as follows :

$$\begin{aligned} V(u)+V(v) &= V(Y_1-Y_2)+2\rho_1\sigma_u\sigma_v, \rho_1 = \rho(u,v) \\ V(u)+V(w) &= V(Y_1-Y_3)+2\rho_2\sigma_u\sigma_w, \rho_2 = \rho(u,w) \\ V(v)+V(w) &= V(Y_2-Y_3)+2\rho_3\sigma_v\sigma_w, \rho_3 = \rho(v,w) \end{aligned}$$

Also we have;

$$\begin{aligned} V(Y_1)-V(Y_2) &= V(u)-V(v) \\ V(Y_1)-V(Y_3) &= V(u)-V(w) \\ V(Y_2)-V(Y_3) &= V(v)-V(w) \end{aligned}$$

From the above systems, we obtain

$$\begin{aligned} V(u)-\rho_1\sigma_u\sigma_v &= \frac{V(Y_1-Y_2)+V(Y_1)-V(Y_2)}{2} = \frac{A}{2} \\ V(u)-\rho_2\sigma_u\sigma_w &= \frac{V(Y_1-Y_3)+V(Y_1)-V(Y_3)}{2} = \frac{B}{2} \end{aligned}$$

$$
\begin{aligned}
V(v) - \rho_1\sigma_u\sigma_v &= \frac{V(Y_1 - Y_2) + V(Y_2) - V(Y_1)}{2} = \frac{C}{2} \\
V(v) - \rho_3\sigma_v\sigma_w &= \frac{V(Y_2 - Y_3) + V(Y_2) - V(Y_3)}{2} = \frac{D}{2} \\
V(w) - \rho_2\sigma_u\sigma_w &= \frac{V(Y_1 - Y_3) + V(Y_3) - V(Y_1)}{2} = \frac{E}{2} \\
V(w) - \rho_3\sigma_v\sigma_w &= \frac{V(Y_2 - Y_3) - V(Y_2) - V(Y_3)}{2} = \frac{F}{2}\,.
\end{aligned}
$$

It may be noted that the differences of error variances are independent of correlation coefficient ρ, as $V(u) - V(v) = \frac{A-C}{2}$, $V(u) - V(w) = \frac{B-E}{2}$, $V(v) - V(w) = \frac{D-F}{2}$. As in the earlier case of independence, have also $V(u) \leq V(v) \leq V(w)$ implies $V(Y_1) \leq V(Y_2) \leq V(Y_2)$. It also follows that once we have a solution for $V(u)$, we can obtain solution for $V(v)$ or $V(w)$ and vice versa. We have,

$$
C\,V(u) - C\rho_1\sigma_u\sigma_v = \frac{AC}{2} \quad \text{and} \quad A\,V(v) - A\rho_1\sigma_u\sigma_v = \frac{AC}{2}.
$$

Subtracting, we get

$$
C\,V(u) - A\,V(w) + (A - C)\,\rho_1\sigma_u\sigma_v = 0.
$$

Solving this equation for $V(v)$, we obtain

$$
V(v) = \frac{C^2}{2C + (A - C)\,\rho_1^2 + \rho_1\sqrt{4AC + (A - C)^2\,\rho_1^2}}.
$$

Again, $F\,V(v) - F\rho_3\sigma_v\sigma_w = \frac{DF}{2}$, $D\,V(w) - D\rho_3\sigma_v\sigma_w = \frac{DF}{2}$, giving $F\,V(v) - DV(w) + (D - F)\,\rho_3\sigma_v\sigma_w = 0$. Solution of this equation gives

$$
V(w) = \frac{F^2}{2F + (D - F)\,\rho_3^2 + \rho_3\sqrt{4DF + (D - F)^2\,\rho_3^2}}.
$$

$V(v)$ and $V(w)$ still contains the unknown parameters ρ_1 and ρ_3. Note that the expressions under square-root signs can not be negative. This allows us to estimate the minimum value of ρ, which allows a real solution for $V(v)$ and $V(w)$. Thus we observe that the above formulation provides us an indirect way to obtain some information about the minimum values of correlation coefficients. The minimum values of correlation coefficients follows from

$$
\rho_1^2 = -\frac{4AC}{(A - C)^2} \quad \text{and} \quad \rho_3^2 = -\frac{4DF}{(D - F)^2}.
$$

Solutions for minimum and ρ_1^2 and ρ_3^2 always exist, since for $\rho_1^2 = 1$ and $\rho_3^2 = 1$, the expressions $4DF + (D-F)^2\,\rho_3^2$ and $4AC + (A-C)^2\,\rho_1^2$ are necessarily positive : reducing to $(D+F)^2$ and $(A+C)^2$ respectively. We may estimate the corresponding error variances for alternative value of ρ_s, ranging from the estimated minimum to the maximum of unity.

5.4.3 Some Discussion on Testing

In this section we outline the possible testing procedure for the error variances obtained through the Naive Approach. If is worth noting that the same formula which are used to solve for population error variances σ_u^2, σ_v^2 and σ_w^2, yield unbiased estimates for the sample values s_u^2, s_v^2 and s_w^2, if we replace the population variances $\sigma_{Y_1}^2, \sigma_{Y_2}^2$ and $\sigma_{Y_3}^2$ by their respective unbiased estimators. Furthermore, the use of F-test to study the significance of difference between s_u^2, s_v^2 and s_w^2 will be biased if basic assumption of independent variates of u, v and w is violated here. If we are not interested in the magnitude of error variances but only in the fact that they are statistically different from each other or not, then we may apply the test procedure disuse by Morgan (1940) and Young (1971).

Morgan (1940) proved that testing $H_0 : \sigma_{y_i}^2, = \lambda^2\ \sigma_{y_j}^2$ against $H_1 : \sigma_{y_i}^2, \neq \lambda^2\ \sigma_{y_j}^2$ is equivalent to testing $H_0 : \rho_{ab} = 0$ against $H_1 : \rho_{ab} \neq 0$, where

$$a = y_i + \lambda y_j, \quad b = y_i - \lambda y, \ \ i \neq j \ \text{ and } \ \lambda = \frac{\sigma y_i}{\sigma y_j}.$$

In the simplest case we may put $\lambda = 1$. This particular test may be applied in our case as it has been mentioned that $\sigma_u^2 \leq \sigma_v^2 \leq \sigma_w^2$ implies $\sigma_{y_1}^2 \leq \sigma_{y_2}^2 \leq \sigma_{y_3}^2$. Testing $H_0 : \sigma_u^2 = \sigma_v^2$ etc. is equivalent to testing $H_0 : \rho_{ab} = 0$ etc.

5.5 Modeling Individual Differences Among Tasters Using Random Effects VC Models

In this section we model the individual differences among assessors using RE linear VC models. For the repeated measurement studies the basic approach discussed in the literature are obviously the repeated measurements ANOVA and MANOVA techniques. We shall investigate the simple one way classification models here with unequal error variances. The heteroscedastic ANOVA models are nothing new in the statistical literature.

For fixed effects models, the estimators of location parameters and the variance components are discussed by Rao et al. (1982), Chen et al. (2000), among many others. The testing procedures for equality of mean effects under heterogeneity are also discussed.

For a particular quality attribute at a time, let y_{ijl} represents the l^{th} replicate of a score given by i^{th} Taster for the sample j, where $i = 1, 2, \ldots, r,\ \ j = 1, 2, \ldots, n$ *and* $l = 1, 2, \ldots, L_{ij}$. For simplicity, we may assume $L_{ij} = L$ for all i and j. This implies that we have typically a balanced data setup where the number of replicates are equal on each sample by each Taster. Also, we may assume that the replications are introduced in a randomized sequence of assessment so that block effects are not necessary in the model. The obvious approach in this situation would be to consider the model

$$y_{ijl} = \mu + \alpha_i + \lambda_j + u_{ijl}\ , \tag{5.8}$$

where u_{ijl} are independent random variates. This is the usual two-way model where λ_js represent the sample specific effects with respect to the particular quality attribute under consideration. α_is correspond to 'average level of assessment' for the Tasters.

Brockhoff et al. (1994) discusses the following parametric formulation which takes scale differences among assessors into account as well as reproducibility differences. Under the assumption of same unknown sample effects λ_j, the model is

$$y_{ijl} = \alpha_i + \beta_i\, \lambda_j + u_{ijl} \tag{5.9}$$

$$Var(u_{ijl}) = \sigma_i^2\ ,\ \sum_j \lambda_j = 0\ ,\ MSE_\lambda = 1.$$

The restrictions over λ_j ensure that the model (5.9) is uniquely parameterized by the space $(\alpha_i\ ,\ \beta_i\ ,\ \sigma_i^2)$ and the λ_js whenever atleast two samples differ. (5.9) is clearly a '*sample* $\times$ *assessor*' interaction model with different assessor variances.

We note here that there is one identification problem associated with the model. If all the sample effects are same, that is, all λ_js are identical, the parameters α_i *and* β_j are not identifiable. This is quite likely when the samples are believed to be the true representatives of a particular population and are likely to posses the same intrinsic property.

In the absence of replication (average of replicated scores may be available) we may study the individual differences among Tasters using the simple one-way heteroscedastic model. In one-way model we simply assume the Taster specific effects as the only

assignable source of variation. If there is sufficient ground to believe that the samples are truly random samples having the same intrinsic characteristic, then the formulation of one-way VC model seems logical. For example, the tea samples (Data Set 5) are essentially the same CTC teas collected from the same experimental garden at the same time and received the same manufacturing process. All the samples were processed in the same factory in the same day. These tea samples may be considered to be truly random samples from the same population of CTC tea. However, the sample variation may be incorporated in the model as an assignable source of variation, leading to a two-way model, and the performance of the two models for the given set of data may always be compared.

For a particular quality attribute, the one-way model may be written as

$$y_{ijk} = \mu_k + \alpha_{ik} + u_{ijk}\,, \tag{5.10}$$

where μ_k represents the overall mean for the attribute k and α_{ik} represents the deviation in i^{th} Taster's average from the overall mean. α_{ik} *and* u_{ijk} are assumed to be independent. In matrix notation the model may be presented as

$$\begin{pmatrix} \underset{\sim 1}{y} \\ \underset{\sim 2}{y} \\ \dots \\ \dots \\ \underset{\sim r}{y} \end{pmatrix} = \mu(1_r \otimes 1_n) + \begin{bmatrix} 1_n & 0 & \dots & \dots & 0 \\ 0 & 1_n & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 1_n \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \dots \\ \alpha_r \end{bmatrix} + \begin{bmatrix} \underset{\sim 1}{u} \\ \underset{\sim 2}{u} \\ \dots \\ \dots \\ \underset{\sim r}{u} \end{bmatrix}$$

$$\Rightarrow y = \mu 1_{nr} + Z\alpha + u\,, \tag{5.11}$$

where $Z = I_r \otimes 1_n$ is the incidence matrix associated with α. The Taster specific effects $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_r)'$ may be assumed to be random. The two possible heteroscedastic formulation of the RE model (5.10) may be proposed. On the basis of two different distributional pattern we distinguish (5.11) by Model 1 and Model 2.

Model 1 : $\alpha_i \sim iid\,(0, \sigma_\alpha^2)$, $u_{ij} \sim iid\,(0, \sigma_i^2)$ and $E(\alpha_i\, u_{ij}) = 0$.

In this formulation we assume that the variation between average level of Tasters' effects is constant and is equal to σ_α^2. But, the variation among Tasters over samples are unequal.

In this case $\sigma_\alpha^2 = 0$ would imply that all the Tasters on average agree on the characteristic of the given set of samples, so far the particular quality attribute is concerned.

Model 2 : $\alpha_i \sim ind\ (0, \sigma_i^2)$, $u_{ij} \sim ind\ (0, \sigma_u^2)$ and $E(\alpha_i\ u_{ij}) = 0$.

Here we assume that the variation between the average level of Tasters' effects is not equal. However, the overall variation among the Tasters over samples is constant (σ_u^2). Here we may test if $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_r^2$, *or* $\sigma_i^2 = 0$ *for all* i.

In the next two sections we study the ANOVA and ML estimators of variance components for the two formulations.

5.6 Estimation of Variance Components For Model 1

5.6.1 ANOVA Estimation

The total sum of squares (TSS) for the model (5.10) may be decomposed as

$$\begin{aligned} \textstyle\sum_i \sum_j (y_{ij} - y_{00})^2 &= \textstyle\sum_i \sum_j (y_{i0} - y_{00})^2 + \sum_i \sum_j (y_{ij} - y_{i0})^2 \\ \Rightarrow TSS &= SS\alpha + SSE \end{aligned}$$

where

$$\begin{aligned} y_{i0} &= \tfrac{1}{n} \textstyle\sum_j y_{ij}\ ,\ y_{00} = \tfrac{1}{r} \sum_i y_{i0}. \\ E(SSE) &= \textstyle\sum_i \sum_j E(u_{ij} - u_{i0})^2 \\ &= \textstyle\sum_i \sum_j [E(u_{ij}^2) + E(u_{i0}^2) - 2E(u_{ij}\ u_{i0})] \qquad (5.12) \\ &= \textstyle\sum_i \sum_j [\sigma_i^2 + \frac{1}{n^2} E(\sum_j u_{ij})^2 - \frac{2}{n} E(u_{ij} \sum_j u_{ij})] \\ &= \textstyle\sum_i \sum_j [\sigma_i^2 + \frac{1}{n} \sigma_i^2 - \frac{2}{n} \sigma_i^2] = (n-1) \sum_i \sigma_i^2 \end{aligned}$$

$$\begin{aligned} E(SSA) &= \textstyle\sum_i \sum_j E[(\alpha_i - \alpha_0) + (u_{i0} - u_{00})]^2 \\ &= \textstyle\sum_i \sum_j [Var(\alpha_i - \alpha_0) + Var(u_{i0} - u_{00})] \\ Var(\alpha_i - \alpha_0) &= Var(\alpha_i) + Var(\alpha_0) - 2\ Cov(\alpha_i, \alpha_0) \\ &= \sigma_\alpha^2 + \tfrac{1}{r} \sigma_\alpha^2 - 2E[\alpha_i \tfrac{1}{r} \textstyle\sum_i \alpha_i] \\ &= \sigma_\alpha^2 + \tfrac{1}{r} \sigma_\alpha^2 - \tfrac{2}{r} \sigma_\alpha^2 = (1 - \tfrac{1}{r}) \sigma_\alpha^2 \\ Var(u_{i0} - u_{00}) &= Var(u_{i0}) + Var(u_{00}) - 2\ Cov(u_{i0}, u_{00}) \\ Var(u_{i0}) &= Var(\tfrac{1}{n} \textstyle\sum_j u_{ij}) = \tfrac{1}{n} \sigma_i^2 \\ Var(u_{00}) &= Var(\tfrac{1}{r} \textstyle\sum_i u_{i0}) = \frac{1}{n\,r^2} \sum_i \sigma_i^2 \end{aligned}$$

$$
\begin{aligned}
Cov(u_{i0}, u_{00}) &= E(u_{i0}\, u_{00}) - E(u_{i0})\, E(u_{00}) \\
&= E[\tfrac{1}{n} \textstyle\sum_j u_{ij}\, \tfrac{1}{nr} \sum_k \sum_l u_{kl}] + 0 \\
&= \tfrac{1}{r\,n^2} E[\textstyle\sum_j u_{ij} \sum_{k \neq i} u_{kl}] + \tfrac{1}{r\,n^2} E[\textstyle\sum_j u_{ij} \sum_l u_{kl}] \\
&= 0 + \tfrac{1}{r\,n^2} E[\textstyle\sum_j u_{ij}\, (u_{i1} + u_{i2} + \ldots + u_{in})] \\
&= \tfrac{1}{r\,n^2} \textstyle\sum_j E(u_{ij}^2) = \tfrac{1}{nr}\, \sigma_i^2 \\
E(SS\alpha) &= n(r-1)\, \sigma_\alpha^2 + (1 - \tfrac{1}{r})\, \textstyle\sum_i \sigma_i^2]
\end{aligned}
\tag{5.13}
$$

From (5.12), we have

$$
\begin{aligned}
& E(SSE) = SSE \\
\Rightarrow\ & (n-1) \sum_i \sigma_i^2 = SSE \\
\Rightarrow\ & \sum_i \sigma_i^2 = \frac{1}{(n-1)}\, SSE
\end{aligned}
$$

Also,

$$
\begin{aligned}
& E(SS\alpha) = SS\alpha \\
\Rightarrow\ & (r-1)\, n\, \sigma_\alpha^2 + \frac{1}{n-1}(1 - \frac{1}{r})\, SSE = SS\alpha
\end{aligned}
$$

Thus, the ANOVA estimator of σ_α^2 is

$$
\hat{\sigma}_\alpha^2 = \frac{1}{n}[\frac{1}{r-1}\, SS\alpha - \frac{1}{r(n-1)}\, SSE]. \tag{5.14}
$$

Again for fixed i, we have

$$
\begin{aligned}
E[\textstyle\sum_j (y_{ij} - y_{i0})^2] &= E[\textstyle\sum_j (u_{ij} - u_{i0})^2] = \textstyle\sum_j [\sigma_i^2 + \tfrac{1}{n}\sigma_i^2 - \tfrac{2}{n}\sigma_i^2] \\
&\Rightarrow\ \sigma_i^2 = \tfrac{1}{n-1} \textstyle\sum_j (y_{ij} - y_{i0})^2\ ,
\end{aligned}
\tag{5.15}
$$

which is the ANOVA type estimator of σ_i^2.

The $\hat{\sigma}_\alpha^2$ and $\hat{\sigma}_i^2$ are unbiased, since

$$
E(\hat{\sigma}_i^2) = \frac{1}{r(n-1)}\, E(SSE) = \sigma_i^2, \quad \text{and}
$$

$$
\begin{aligned}
E(\hat{\sigma}_\alpha^2) &= \frac{1}{n(r-1)} E(SS\alpha) - \frac{1}{nr(n-1)} E(SSE) \\
&= \frac{1}{n(r-1)}[n(r-1)\sigma_\alpha^2 + \frac{1}{r}(r-1)\sum_i \sigma_i^2] - \frac{1}{nr(n-1)}(n-1)\sum_i \sigma_i^2 \\
&= \sigma_\alpha^2 + \frac{1}{nr}\sum_i \sigma_i^2 - \frac{1}{nr}\sum_i \sigma_i^2 = \sigma_\alpha^2.
\end{aligned}
$$

To establish the minimum variance property we may proceed as follows :

$$
\begin{aligned}
y'y &= SS\mu + SS\alpha + SSE \\
&= \tfrac{1}{n\,r}\, y_{00}^2 + \textstyle\sum_i \sum_j (y_{i0} - y_{00})^2 + \sum_i \sum_j (y_{ij} - y_{i0})^2
\end{aligned}
$$

We have, $y_{i0} = \mu + \alpha_i + u_{i0} \sim N(\mu, \sigma_\alpha^2 + \frac{1}{n}\sigma_i^2)$ under the prior assumption of normality of α_i and u_{ij}. If we denote $y_0' = (\underset{\sim}{y}_{10}, \underset{\sim}{y}_{20} \ldots \underset{\sim}{y}_{r0})$ then $\underset{\sim}{y}_0' \sim N(1_r\mu, (\sigma_\alpha^2 + \frac{1}{n}\sigma_i^2)I_r)$.

Thus $SS(\alpha) = \sum_i \sum_j (y_{i0} - y_{00})^2 = \underset{\sim}{y}_0'[I_r - \frac{1}{r}1_r 1_r'] \underset{\sim}{y}_0$ follows $(\sigma_\alpha^2 + \frac{1}{n}\sigma_i^2)\,\chi_{r-1}^2$. Now for SSE, we note that $y_{ij} - y_{i0} = u_{ij} - u_{i0}$. Since $u_{ij} \sim N(0, \sigma_i^2)$ and u_{ij} are all mutually independent, we can write for each i.

$$
a_i = \sum_{j=1}(y_{ij} - y_{i0})^2 = \sum_j (u_{ij} - u_{i0})^2 \sim \sigma_i^2\, \chi_{n-1}^2
$$

and a_i $(i = 1, 2, \ldots n)$ are mutually independent. Thus $SSE = \sum_{i=1}^r a_i^2 \sim \sigma_i^2\, \chi_{r(n-1)}^2$.

Finally we are to show that $SS\alpha$ and SSE are independent. For the identity

$$
y_{ij} = y_{00} + (y_{i0} - y_{00}) + (y_{ij} - y_{i0}) = a + b_i + c_{ij},
$$

it is sufficient to show that b_i and c_{ij} are independent. Since b_i and c_{ij} are normally distributed, it is sufficient to show that they are uncorrelated. It means that for all i, i' and j, $Cov(b_i, c_{i'j}) = 0$, which may easily be shown. Hence the ANOVA estimators of variance components of the heteroscedastic one-way model are MVUE.

5.6.2 ML Estimation

The variance of composite random component $e_{ij} = \alpha_i + u_{ij}$ of the model (5.10) is $\sigma_\alpha^2 + \sigma_i^2$. The variance-covariance matrix may be written as

$$
\begin{aligned}
E(e\,e') = \Phi_1 &= Z\, E(\alpha\alpha')\, Z' + E(u\,u') \\
&= Z\sigma^2 I_r Z' + \Sigma_i \otimes I_n = \sigma_\alpha^2(I_r \otimes J_n) + \Sigma_i \otimes I_n\,,
\end{aligned}
$$

where $Z\,Z' = I_r \otimes J_n$ and $\Sigma_i = diag\ (\sigma_i^2,\ \sigma_i^2 \ldots \sigma_i^2)$. We may write $\Phi_1 = diag(S_1, S_2, \ldots, S_r)$, where

$$S_i = \begin{bmatrix} \sigma_\alpha^2 + \sigma_i^2 & \sigma_\alpha^2 & \ldots & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\alpha^2 + \sigma_i^2 & \ldots & \sigma_\alpha^2 \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \ldots & \sigma_\alpha^2 + \sigma_i^2 \end{bmatrix}, \quad i = 1, 2, \ldots r.$$

The determinant and inverse of Φ_1 may be obtained as follows :

$$| S_i | = \begin{bmatrix} \sigma_\alpha^2 + \sigma_i^2 & \sigma_\alpha^2 & \ldots & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\alpha^2 + \sigma_i^2 & \ldots & \sigma_\alpha^2 \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \ldots & \sigma_\alpha^2 + \sigma_i^2 \end{bmatrix} = \begin{bmatrix} \sigma_\alpha^2 + n\sigma_i^2 & \sigma_\alpha^2 + n\sigma_i^2 & \ldots & \sigma_\alpha^2 + n\sigma_i^2 \\ \sigma_\alpha^2 & \sigma_\alpha^2 + \sigma_i^2 & \ldots & \sigma_\alpha^2 \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \ldots & \sigma_\alpha^2 + \sigma_i^2 \end{bmatrix}$$

$$= \frac{\sigma_i^2 + n\sigma_\alpha^2}{\sigma_\alpha^2} \begin{bmatrix} \sigma_\alpha^2 & \sigma_\alpha^2 & \ldots & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\alpha^2 + \sigma_i^2 & \ldots & \sigma_\alpha^2 \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \ldots & \sigma_\alpha^2 + \sigma_i^2 \end{bmatrix} = \frac{\sigma_i^2 + n\sigma_\alpha^2}{\sigma_\alpha^2} \begin{bmatrix} \sigma_\alpha^2 & \sigma_\alpha^2 & \ldots & \sigma_\alpha^2 \\ 0 & \sigma_i^2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \sigma_i^2 \end{bmatrix}$$

$$= (\sigma_i^2 + n\,\sigma^2\alpha)\ (\sigma_i^2)^{n-1}$$

Thus,

$$| \Phi_1 | = | S_1 |\ |\ S_2 | \ldots | S_r | = \left[\prod_i (\sigma_i^2 + n\sigma_\alpha^2)\right]\ \left[\prod_i \sigma_i^2\right]^{n-1} \tag{5.16}$$

Also,

$$
\begin{aligned}
S_i^{-1} &= (\sigma_i^2\, I_n + \sigma_\alpha^2\, J_n)^{-1} = (\sigma_\alpha^2)^{-1}\left[1_n 1_n' + \tfrac{\sigma_i^2}{\sigma_\alpha^2}\, I_n\right]^{-1} \\
&= (\sigma_\alpha^2)^{-1}\left[\left(\tfrac{\sigma_i^2}{\sigma_\alpha^2}\, I_n\right)^{-1} - \left(\frac{[(\frac{\sigma_i^2}{\sigma_\alpha^2})^{-1} I_n)^{-1} 1_n]\,[1_n'(\frac{\sigma_i^2}{\sigma_\alpha^2} I_n)^{-1}]}{1+1_n'(\frac{\sigma_i^2}{\sigma_\alpha^2} I_n)^{-1} 1_n}\right)\right] \\
&= \tfrac{1}{\sigma_i^2}\left[I_n - \frac{\frac{\sigma_\alpha^2}{\sigma_i^2} I_n 1_n 1_n' I_n}{1+\frac{\sigma_\alpha^2}{\sigma_i^2} 1_n' I_n 1_n}\right] \\
&= \tfrac{1}{\sigma_i^2}\left[I_n - \tfrac{\sigma_\alpha^2}{\sigma_i^2+n\sigma_\alpha^2}\, J_n\right]
\end{aligned}
\tag{5.17}
$$

In finding the S_i^{-1} we have used the formula presented in the Rao's book (Rao, 1973, p. 33).

Under normality assumption, the probability function of the response y may be written as

$$p(y) = (2\pi)^{-\frac{nr}{2}} \mid \Phi_1 \mid^{-\frac{1}{2}}\ exp[-\frac{1}{2}\,(y-\mu\, 1_{nr})'\,\Phi_1^{-1}(y-\mu\, 1_{nr})].$$

The likelihood function for Model 1 may be written as

$$L_1 = (2\pi)^{-\frac{nr}{2}}\,(\Pi_i\,\sigma_i^2)^{\frac{n-1}{2}}\left[\Pi_i\,(\sigma_i^2+n\sigma_\alpha^2)\right]^{-\frac{1}{2}}\ exp\ \left[-\frac{1}{2}(y-\mu\, 1_{nr})'\,\Phi_1^{-1}(y-\mu\, 1_{nr})\right], \tag{18}$$

and the quadratic form in the exponent may be written as

$$
\begin{aligned}
&(y-\mu 1_{nr})'\,\Phi_1^{-1}(y-\mu 1_{nr}) \\
&= \sum_{i=1}^{r}(\underset{\sim i}{y}-\mu 1_n)'\,\frac{1}{\sigma_i^2}\left[I_n - \frac{\sigma_\alpha^2}{\sigma_i^2+n\sigma_\alpha^2}J_n\right](\underset{\sim i}{y}-\mu 1_n) \\
&= \sum_i \frac{1}{\sigma_i^2}\sum_j (y_{ij}-\mu)^2 - \sum_i \frac{1}{\sigma_i^2}\frac{\sigma_\alpha^2}{\sigma_i^2+n\sigma_\alpha^2}[\sum_j (y_{ij}-\mu)]^2\ .
\end{aligned}
$$

The loglikelihood function would be

$$
\begin{aligned}
l_1 &= -\tfrac{nr}{2}\,\ln(2\pi) - \tfrac{n-1}{2}\textstyle\sum_i \ln\sigma_i^2 - \tfrac{1}{2}\,\sum_i \ln(\sigma_i^2+n\sigma_\alpha^2) \\
&\quad -\tfrac{1}{2}\,\textstyle\sum_i\sum_j \tfrac{1}{\sigma_i^2}(y_{ij}-\mu)^2 + \tfrac{1}{2}\,\sum_i \tfrac{1}{\sigma_i^2}\tfrac{\sigma_\alpha^2}{\sigma_i^2+n\sigma_\alpha^2}\,[\textstyle\sum_j (y_{ij}-\mu)]^2
\end{aligned}
\tag{5.19}
$$

Partially differentiating (5.19) with respect to μ, σ_α^2 and σ_i^2 respectively, the ML estimators of the location parameters and the variance components may be obtained from the following relations :

$$\begin{aligned} \frac{\partial l_1}{\partial \mu} = 0 \;&\Rightarrow\; \textstyle\sum_i \sum_j \frac{1}{\sigma_i^2}\left[1 - \frac{n\sigma_\alpha^2}{\sigma_i^2+n\sigma_\alpha^2}\right](y_{ij} - \mu) = 0 \\ &\Rightarrow\; \textstyle\sum_i \sum_j\; w_i\, y_{ij} / \sum_i w_i = n\,\mu \\ &\Rightarrow\; \mu = \tfrac{1}{n}\; \textstyle\sum_j\; y_j^*. \end{aligned} \tag{5.20}$$

where $w_i = \frac{1}{\sigma_i^2+n\sigma_\alpha^2}$ is the weight. Clearly ML estimator of μ is the weighted mean where w_i is the weight for i^{th} response. Note that the GLSE of μ may be shown to be of the form : GLSE $(\mu) = \frac{\sum_i y_{i0}/Var(y_{i0})}{\sum_i \frac{1}{Var(y_{i0})}}$, where $Var(y_{i0}) = Var(\mu+\alpha_i+u_{i0}) = \sigma_\alpha^2 + \frac{1}{n}\,\sigma_i^2$, which implies that $w_i = \frac{1}{Var(y_{i0})}$. Also,

$$\frac{\partial l_1}{\partial \sigma_\alpha^2} = 0 \Rightarrow -\sum_i \frac{1}{\sigma_i^2 + n\sigma_\alpha^2}\left[1 - \frac{n}{\sigma_i^2 + n\,\sigma_\alpha^2}\right](y_{i0} - \mu)^2 = 0\,, \tag{5.21}$$

$$\frac{\partial l_1}{\partial \sigma_i^2} = 0 \Rightarrow \frac{\sigma_i^2}{\sigma_i^2 + n\sigma_\alpha^2} + \frac{\sigma_\alpha^2}{\sigma_i^2 + n\sigma_\alpha^2}\left[\frac{1}{\sigma_i^2} + \frac{1}{\sigma_i^2 + n\sigma_\alpha^2}\right] D_i^2 = \frac{1}{\sigma_i^2}\sum_j (y_{ij} - \mu)^2 - (n-1). \tag{5.22}$$

where $D_i = \sum_j (y_{ij} - \mu)$. Note that for unequal σ_i^2s, explicit expressions for σ_i^2 and σ_α^2 can not be obtained from (5.21) and (5.22).

5.6.3 Estimation of Random Component of the Model

We briefly discuss the prediction aspects in the heteroscedastic one-way model (5.10) in this section. We start with the prediction of the random component α_i. The general theory of prediction for the random/mixed effects models is discussed by Searle et al. (1992).

We may think of some related information, say y_{i0}, for the unobservable random component α_i. Now the question is, can we think of some numerical value of α_i (say α_i^*) on the basis of y_{i0}? In predicting α_i, it is sensible to consider $E(\alpha_i)$ as the predictor. That is, α_i^* may be taken as $E(\alpha_i)$; but $E(\alpha_i) = 0$. Again if we can think that y_{i0} is considerably larger than the overall average μ, then we may expect that α_i is positive ($y_{i0} = \mu + \alpha_i + u_{i0}$ and $y_{i0} > \mu$ implies $\alpha_i > 0$). With this thought, we may use the conditional mean $E(\alpha_i/y_{i0})$ rather than $E(\alpha_i)$ as our predictor.

Note that α_i and y_{i0} jointly follows bivariate normal distribution with

$$E\begin{bmatrix} \alpha_i \\ y_{i0} \end{bmatrix} = \begin{bmatrix} 0 \\ \mu \end{bmatrix} \quad \text{and} \quad Var\begin{bmatrix} \alpha_i \\ y_{i0} \end{bmatrix} = \begin{bmatrix} \sigma_\alpha^2 & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\alpha^2 + \frac{1}{n}\sigma_i^2 \end{bmatrix}.$$

Thus, we have

$$\begin{aligned} E(\alpha_i/y_{i0}) &= E(\alpha_i) + Cov(\alpha_i, y_{i0})\,[Var(y_{i0})]^{-1}(y_{i0} - \mu) \\ &= \sigma_\alpha^2\,(\sigma_\alpha^2 + \tfrac{1}{n}\sigma_i^2)^{-1}(y_{i0} - \mu) = \tfrac{n\,\sigma_\alpha^2}{n\,\sigma_\alpha^2 + \sigma_i^2}\,(y_{i0} - \mu)\,, \end{aligned} \tag{5.23}$$

which is the predictor of α_i.

An alternative approach of estimating α_i would be to consider the conditional distribution of α_i given the total errors $\alpha_i + u_{i1}, \alpha_i + u_{i2}, \ldots, \alpha_i + u_{in}$. An appropriate summary measure of this conditional distribution may be assumed to represent the deviation due to i^{th} Taster's effect. The conditional distribution being normal, the summary measure would be the same for mean, median or mode.

We have, $\underset{\sim}{\alpha}^* \sim N(0, \Omega)$, where

$$\alpha^* = \begin{bmatrix} \alpha_i + u_{i1} \\ \alpha_i + u_{i2} \\ \cdots \\ \cdots \\ \alpha_i + u_{in} \\ \alpha_i \end{bmatrix} = \begin{bmatrix} y_{i1} - \mu \\ y_{i2} - \mu \\ \cdots \\ \cdots \\ y_{in} - \mu \\ y_i \end{bmatrix} \text{ and } \Omega = \begin{bmatrix} \sigma_\alpha^2 + \sigma_i^2 & \sigma_\alpha^2 & \ldots & \ldots & \sigma_\alpha^2 & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\alpha^2 + \sigma_i^2 & \ldots & \ldots & \sigma_\alpha^2 & \sigma_\alpha^2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \ldots & \ldots & \sigma_\alpha^2 + \sigma_i^2 & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \ldots & \ldots & \sigma_\alpha^2 & \sigma_\alpha^2 \end{bmatrix}$$

The conditional probability of α_i given the total errors is

$$\begin{aligned} p(.) &= p(\alpha_i \mid \alpha_i + u_{i1}, \ldots \alpha_i + u_{in}) = \frac{p(\alpha_i + u_{i1}, \ldots \alpha_i + u_{in}, \alpha_i)}{p(\alpha_i + u_{i1}, \ldots \alpha_i + u_{in})} \\ &= (\sqrt{2\pi})^{-(n+1)}\ |\,\Omega\,|^{-\frac{1}{2}}\ exp\left[-\frac{1}{2}\alpha^{*\prime}\,\Omega^{-1}\,\alpha^*\right]/L^* = \frac{N}{L^*}\,, \end{aligned}$$

where

$$L^* = (\sqrt{2\pi})^{-n}\ |\,S_i\,|^{-\frac{1}{2}}\ exp\left[-\frac{1}{2}(y_i - \mu)'\,S_i^{-1}(y_i - \mu)\right]$$

$$= (2\pi)^{-\frac{n}{2}}(\sigma_i^2)^{-\frac{n-1}{2}}\ (\sigma_i^2+\sigma_\alpha^2)^{-\frac{1}{2}}$$
$$exp\left[-\frac{1}{2\sigma_i^2}\left\{\sum_j(y_{ij}-\mu)^2-\frac{\sigma_\alpha^2}{\sigma_i^2+n\sigma_\alpha^2}\left\{\sum_j(y_{ij}-\mu)\right\}^2\right\}\right]$$

To obtain Ω^{-1} we use the following result (Rao, 1973, p. 33) :

$$\begin{pmatrix} S_i & B \\ B' & D \end{pmatrix}^{-1} = \begin{bmatrix} S_i^{-1}+FE^{-1}F' & -FE^{-1} \\ -E^{-1}F' & E^{-1} \end{bmatrix},$$

$E = D - B'\ S_i^{-1}\ B$, $F = S_i^{-1}B$, where $B' = \sigma_\alpha^2\ \ 1_n'$ and S_i is defined in Section 5.5.2. We obtain

$$E = D - \sigma_\alpha^2\ 1_n'\left[\frac{1}{\sigma_i^2}I_n - \frac{\sigma_\alpha^2}{\sigma_i^2(\sigma_i^2+n\sigma_\alpha^2)}\ J_n\right]\sigma_\alpha^2\ 1_n$$

$$= D - \frac{n\sigma_\alpha^4}{\sigma_i^2+n\sigma_\alpha^2} = \frac{\sigma_\alpha^2\ \sigma_i^2}{\sigma_i^2+n\sigma_\alpha^2}\ ,\ \text{for}\ \ D=\sigma_\alpha^2.$$

$$F = S_i^{-1}B = \left[\frac{1}{\sigma_i^2}I_n - \frac{\sigma_\alpha^2}{\sigma_i^2(\sigma_i^2+n\sigma_\alpha^2)}\ J_n\right]\sigma_\alpha^2\ 1_n = \frac{\sigma_\alpha^2}{\sigma_i^2+n\sigma_\alpha^2}\ 1_n$$

$$FE^{-1}F' = \frac{\sigma_\alpha^2}{\sigma_i^2(\sigma_i^2+n\sigma_\alpha^2)}\ J_n\ ,\ \text{and}\ \ S_i^{-1}+FE^{-1}F' = \frac{1}{\sigma_i^2}\ I_n$$

$$E^{-1}F' = \frac{1}{\sigma_i^2}1_n = FE^{-1}\ .$$

Thus we may write,

$$\Omega^{-1} = \frac{1}{\sigma_i^2}\begin{bmatrix} I_n & -1_n \\ 1_n' & \frac{\sigma_i^2+n\,\sigma_\alpha^2}{\sigma_\alpha^2} \end{bmatrix},\ \textit{and}$$
$$|\Omega| = |S_i||D - B'S_i^{-1}B| = \sigma_\alpha^2(\sigma_i^2)^n$$

Thus we may write

$$p(.) = \frac{\sqrt{\sigma_i^2+n\,\sigma_\alpha^2}}{\sqrt{2\pi\sigma_i^2\,\sigma_\alpha^2}}\ exp\left[-\frac{1}{2\sigma_i^2}\left\{(\sum_j y_{ij} - n\,\mu)^2\frac{\sigma_\alpha^2}{\sigma_i^2+n\,\sigma_\alpha^2}\right.\right.$$

$$
\begin{aligned}
& -2\,\alpha_i\,(\sum_j y_{ij} - n\,\mu) + \frac{\alpha_i^2\,(\sigma_i^2 + n\,\sigma_\alpha^2)}{\sigma_\alpha^2} \Bigg\} \Bigg] \\
= \; & \frac{\sqrt{\sigma_i^2 + n\,\sigma_\alpha^2}}{\sqrt{2\pi\sigma_i^2\,\sigma_\alpha^2}}\; exp\left[-\frac{(\sigma_i^2 + n\,\sigma_\alpha^2)}{2\,\sigma_i^2\,\sigma_\alpha^2} \left\{ \alpha_i^2 - 2\frac{\alpha_i\,\sigma_\alpha^2}{\sigma_i^2 + n\,\sigma_\alpha^2}\,(\sum_j y_{ij} - n\,\mu) \right.\right. \\
& \left.\left. +(\sum_j y_{ij} - n\,\mu)^2 \left(\frac{\sigma_\alpha^2}{\sigma_i^2 + n\,\sigma_\alpha^2}\right)^2 \right\} \right] \\
= \; & \frac{\sqrt{\sigma_i^2 + n\,\sigma_\alpha^2}}{\sqrt{2\pi\sigma_i^2\,\sigma_\alpha^2}}\; exp\left[-\frac{(\sigma_i^2 + n\,\sigma_\alpha^2)}{2\,\sigma_\alpha^2\,\sigma_i^2} \left\{ \alpha_i - \frac{\sigma_\alpha^2}{\sigma_i^2 + n\,\sigma_\alpha^2} \left(\sum_j y_{ij} - n\,\mu \right) \right\}^2 \right]
\end{aligned}
$$

Thus, the distribution of α_i may be written as

$$
\alpha_i \sim N\left\{ \frac{n\,\sigma_\alpha^2}{\sigma_i^2 + n\,\sigma_\alpha^2}(\bar{y}_i - \mu)\,,\, \frac{\sigma_i^2\,\sigma_\alpha^2}{\sigma_i^2 + n\,\sigma_\alpha^2} \right\}
$$

The approximate predictor of α_i may thus be written as

$$
\hat{\alpha}_i \;=\; \frac{n\,\hat{\sigma}_\alpha^2}{\hat{\sigma}_i^2 + n\,\hat{\sigma}_\alpha^2}(y_{i0} \;-\; \hat{\mu}) \tag{5.24}
$$

5.7 Estimation of Variance Components for Model 2

5.7.1 ANOVA Estimation

For the Model 2, we assume $\alpha_i \sim iid\ (0, \sigma_i^2)$ and $u_{ij} \sim iid\ (0, \sigma_u^2)$. The partition of TSS is

$$
\begin{aligned}
\sum_i \sum_j (y_{ij} - y_{00})^2 \;&=\; \sum_i \sum_j (y_{i0} - y_{00})^2 + \sum_i \sum_j (y_{ij} - y_{i0})^2 \\
\Rightarrow \; TSS \;&=\; SS\alpha + SSE\,.
\end{aligned}
$$

The y_{i0} and y_{00} have already been defined in Section 5.5.1.

$$
\begin{aligned}
E(SSE) \;&=\; \Sigma_i \Sigma_j \left[E(u_{ij}^2) + E(u_{i0}^2) - 2E(u_{ij}\,u_{i0}) \right] \\
&=\; nr\,\left[\sigma_u^2 + \tfrac{1}{n^2}\,E(\Sigma_j\,u_{ij})^2 - \tfrac{2}{n}\,E(u_{ij}\;\Sigma_j\,u_{ij}) \right] \\
&=\; nr\,\left[\sigma_u^2 + \tfrac{1}{n}\,\sigma_u^2 - \tfrac{2}{n}\sigma_u^2 \right] = r(n-1)\,\sigma_u^2
\end{aligned} \tag{5.25}
$$

$$
\begin{aligned}
E(SS\alpha) \;&=\; \Sigma_i \Sigma_j \left[E(\alpha_i - \alpha_0)^2 + E(u_{i0} - u_{00})^2 - 2E(\alpha_i - \alpha_0)(u_{i0} - u_{00}) \right] \\
&=\; \Sigma_i \Sigma_j \left[Var(\alpha_i - \alpha_0) + Var(u_{i0} - u_{00}) \right]
\end{aligned} \tag{5.26}
$$

$$
\begin{aligned}
Var\,(\alpha_i - \alpha_0) &= Var\,(\alpha_i) + Var\,(\alpha_0) - 2\,Cov\,(\alpha_i, \alpha_0) \\
&= \sigma_i^2 + Var\,(\tfrac{1}{r}\,\textstyle\sum_i\,\alpha_i) - 2E(\alpha_i \tfrac{1}{r}\,\textstyle\sum_i\,\alpha_i) \\
&= \sigma_i^2 + \tfrac{1}{r^2}\,\textstyle\sum_i\,\sigma_i^2 - \tfrac{2}{r}\,\sigma_i^2
\end{aligned}
$$

$$
\begin{aligned}
Var\,(u_{i0} - u_{00}) &= \tfrac{1}{n^2}\,Var\,(\textstyle\sum_j\,u_{ij}) + \tfrac{1}{n^2r^2}\,Var\,(\textstyle\sum_i\,\sum_j\,u_{ij}) - 2E\,(u_{i0}\,u_{00}) \\
&= \tfrac{1}{n}\,\sigma_u^2 + \tfrac{1}{nr}\,\sigma_u^2 - 2E(u_{i0}\,u_{00}) \\
E(u_{i0}\,u_{00}) &= E\left[\tfrac{1}{n}\textstyle\sum_j u_{ij}\,\tfrac{1}{nr}\sum_k\sum_l u_{kl}\right] \\
&= \tfrac{1}{n^2r}E\left[\textstyle\sum_j u_{ij}\;\sum_k\sum_{l\neq j} u_{kl}\right] + \tfrac{1}{n^2r}E\left[\textstyle\sum_j u_{ij}\sum_l u_{kl}\right] \\
&= 0 + \tfrac{1}{n^2r}E\left[\textstyle\sum_j u_{ij}(u_{i1} + u_{i2} + \ldots + u_{in})\right] \\
&= \tfrac{1}{n^2r}E\left[\textstyle\sum_j u_{ij}^2\right] = \tfrac{1}{nr}\,\sigma_u^2 \\
E(u_{i0} - u_{00})^2 &= \tfrac{1}{n}\,\sigma_u^2 + \tfrac{1}{nr}\,\sigma_u^2 - \tfrac{2}{nr}\,\sigma_u^2 = \tfrac{1}{n}\,(1 - \tfrac{1}{r})\,\sigma_u^2
\end{aligned}
$$

Thus from (5.26)

$$
\begin{aligned}
E(SS\alpha) &= \textstyle\sum_i\,\sum_j\,\left[(1 - \frac{2}{r})\,\sigma_i^2 + \frac{1}{r^2}\,\sum_i\,\sigma_i^2 + \frac{1}{n}(1 - \frac{1}{r})\,\sigma_u^2\right] \\
&= n(1 - \tfrac{1}{r})\,\textstyle\sum_i\,\sigma_i^2 + (r-1)\sigma_u^2
\end{aligned}
\qquad (5.27)
$$

From (5.25), equating $E(SSE)$ to SSE, the ANOVA estimator of σ_u^2 is obtained as

$$
\sigma_u^2 = \frac{1}{r(n-1)} SSE\,. \qquad (5.28)
$$

Also, for fixed i, starting with the form $\sum_j (y_{ij} - y_{i0})^2$ and using the relation (5.27), the ANOVA estimator of σ_i^2 may be obtained from the following relation

$$
n(1 - \frac{2}{r})\,\sigma_i^2 = \sum_j (y_{ij} - y_{i0})^2 - (1 - \frac{2}{r})\,\sigma_u^2 - \frac{1}{r(r-1)}\,SS\alpha \qquad (5.29)
$$

These ANOVA estimators may easily be shown to be unbiased. Also they are MVUE. Using the following result given by Rao et al. (1982), the large sample variance of the location parameter & variance components may be calculated.

5.7.2 ML Estimation

The variance of the composite random component $e_{ij} = \alpha_i + u_{ij}$ of the Model 2 is $\sigma_i^2 + \sigma_u^2$. The variance-covariance matrix may be written as

$$E(e\, e') = \Phi_2 = ZE(\alpha\alpha')\, Z' + E(u\, u') = Z\, \Sigma\, Z' + \sigma_u^2\, I_{nr}\ ,$$

where $\Sigma = diag\ (\sigma_1^2, \sigma_2^2, \ldots, \sigma_r^2)$. Also it may be easily shown that $Z\, \Sigma\, Z' = \Sigma \otimes J_n$. Thus we have, $\Phi_2 = diag\ (S_1, S_2, \ldots, S_r)$, where

$$S_i = \begin{bmatrix} \sigma_i^2 + \sigma_u^2 & \sigma_i^2 & \ldots & \sigma_i^2 \\ \sigma_i^2 & \sigma_i^2 + \sigma_u^2 & \ldots & \sigma_i^2 \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \sigma_i^2 & \sigma_i^2 & \ldots & \sigma_i^2 + \sigma_u^2 \end{bmatrix}_{n \times n}$$

The determined and inverse of S_i may be obtained as follows :

$$| S_i | = \begin{bmatrix} \sigma_u^2 + n\sigma_i^2 & \sigma_u^2 + n\sigma_i^2 & \ldots & \sigma_u^2 + n\sigma_i^2 \\ \sigma_i^2 & \sigma_i^2 + \sigma_u^2 & \ldots & \sigma_i^2 \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \sigma_i^2 & \sigma_i^2 & \ldots & \sigma_i^2 + \sigma_u^2 \end{bmatrix} = \frac{\sigma_u^2 + n\sigma_i^2}{\sigma_i^2} \begin{bmatrix} \sigma_i^2 & \sigma_i^2 & \ldots & \sigma_i^2 \\ \sigma_i^2 & \sigma_u^2 + \sigma_i^2 & \ldots & \sigma_i^2 \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \sigma_i^2 & \sigma_i^2 & \ldots & \sigma_i^2 + \sigma_u^2 \end{bmatrix}$$

$$= \frac{\sigma_u^2 + n\sigma_i^2}{\sigma_i^2} \begin{bmatrix} \sigma_i^2 & \sigma_i^2 & \ldots & \sigma_i^2 \\ 0 & \sigma_u^2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \sigma_u^2 \end{bmatrix} = (\sigma_u^2)^{n-1}\, (\sigma_u^2 + n\sigma_i^2), \text{ and thus}$$

$$| \Phi_2 | = (\sigma_u^2)^{r(n-1)}\, \Pi_{i=1}^{r}\, (\sigma_u^2 + n\sigma_i^2) \tag{5.30}$$

Following the procedure adopted in obtaining the form (5.17), the inverse of S_i may easily be obtained as

$$S_i^{-1} = \frac{1}{\sigma_u^2} \left[I_n - \frac{\sigma_i^2}{\sigma_u^2 + n\sigma_i^2}\, J_n \right] \tag{5.31}$$

Under normality assumption, the likelihood function for Model 2 may be written as

$$L_2 = (2\pi)^{-\frac{nr}{2}} \ (\sigma_u^2)^{-\frac{r(n-1)}{2}} \ \left[\Pi_i(\sigma_u^2 + n\sigma_i^2)\right]^{-\frac{1}{2}} \ exp\left[-\frac{1}{2}(y - \mu 1_{nr})' \ \Phi_2^{-1}(y - \mu 1_{nr})\right] \ ,$$

and the quadratic form in the exponent may be simplified as

$$\begin{aligned}
& (y - \mu 1_{nr})' \ \Phi_2^{-1}(y - \mu 1_{nr}) \\
&= \sum_{i=1}^{r}(y_i - \mu 1_n)' \left[\frac{1}{\sigma_u^2} I_n - \frac{\sigma_i^2}{\sigma_u^2 \ (\sigma_u^2 + n\sigma_i^2)} J_n\right] (y - \mu 1_{nr}) \\
&= \frac{1}{\sigma_u^2}\sum_i\sum_j(y_{ij} - \mu)^2 - \frac{1}{\sigma_u^2}\sum_i \frac{\sigma_i^2}{\sigma_u^2 + n\sigma_i^2}\left[\sum_j(y_{ij} - \mu)\right]^2
\end{aligned}$$

The loglikelihood function would be

$$\begin{aligned}
l_2 \ = \ & -\tfrac{nr}{2} \ ln(2\pi) - \tfrac{r(n-1)}{2} \ ln \ \sigma_u^2 - \tfrac{1}{2} \ \textstyle\sum_i \ ln(\sigma_u^2 + n\sigma_i^2) \\
& -\tfrac{1}{2\sigma_u^2} \textstyle\sum_i \sum_j (y_{ij} - \mu)^2 + \tfrac{1}{2\sigma_u^2} \textstyle\sum_i \ \tfrac{\sigma_i^2}{\sigma_u^2 + n\sigma_i^2} \left[\sum_j (y_{ij} - \mu)\right]^2
\end{aligned} \tag{5.32}$$

Partially differentiating (5.32) with respect to μ, σ_u^2 and σ_i^2 respectively, the ML estimators of location parameters and variance components may be obtained from the following relations :

$$\begin{aligned}
\tfrac{\partial l_2}{\partial \mu} = 0 \ \Rightarrow \ & \textstyle\sum_i \ \tfrac{n}{\sigma_u^2 + n\sigma_i^2} \ (y_{ij} - \mu) = 0 \\
\Rightarrow \ & \mu = \tfrac{\sum a_i y_{i0}}{\sum a_i} \ ,
\end{aligned} \tag{5.33}$$

where $a_i = \frac{n}{\sigma_u^2 + n\sigma_i^2}$ is the weight. The GLS estimator of μ is $\mu = \frac{\sum_i y_{i0}/Var(y_{i0})}{\sum_i \frac{1}{Var(y_{i0})}}$, where $Var(y_{i0}) = \sigma_i^2 + \frac{1}{n}\sigma_u^2 = (a_i)^{-1}$. This implies that $a_i = \frac{1}{Var(y_{i0})}$, which further implies that GLS and ML estimators of location parameter are same. Also, it may be easily shown that

$$\sigma_i^2 = (y_{i0} - \mu)^2 - \frac{1}{n}\sigma^2 \ , \tag{5.34}$$

$$\begin{aligned}
& \tfrac{\partial l_2}{\partial \sigma_u^2} = 0 \qquad \Rightarrow \ (n-1)r + \textstyle\sum_i \ \tfrac{\sigma_u^2}{\sigma_u^2 + n\sigma_i^2} + (n-1) \ \textstyle\sum_i \tfrac{S_i^2}{\sigma_i^2} - \\
& \tfrac{1}{\sigma_u^2} \ \textstyle\sum_i \ \tfrac{\sigma_i^2}{\sigma_u^2 + n\sigma_i^2}\left[1 + \tfrac{\sigma_i^2 \sigma_u^2}{\sigma_u^2 + n\sigma_i^2}\right] \ D_i^2 = 0 \\
& \qquad\qquad \Rightarrow \ r(n-1)\textstyle\sum_i \tfrac{\sigma_u^2}{d_i} + (n-1) \ \textstyle\sum_i \ S_i^2/\sigma_i^2 - \tfrac{1}{\sigma_u^2} \ \textstyle\sum_i \ \sigma_i^2/d_i \left[1 + \tfrac{\sigma_i^2 \sigma_u^2}{d_i}\right] \ D_i^2
\end{aligned} \tag{5.35}$$

where $S_i^2 = (n-1)^{-1} \sum_j (y_{ij} - \mu)$, $D_i = \sum_j (y_{ij} - \mu)$ and $d_i = \sigma_u^2 + n\,\sigma_i^2$. As in case of Model 1, for unequal σ_i^2s explicit expressions for σ_i^2 and σ_u^2 can not be obtained from (5.34) and (5.35), for the Model 2 as well.

The approach to estimate (predict) α_i here is exactly same as discussed for Model 1 in Section 5.5.3. The conditional expectation of α_i given y_{i0} may be obtained as

$$E(\alpha_i \mid y_{i0}) = \frac{n\sigma_i^2}{\sigma_u^2 + n\sigma_i^2}(y_{i0} - \mu) \,,$$

which is the predictor of α_i. Again, from the conditional distribution of α_i given the total errors, the appropriate measure of α_i may be obtained easily following the procedure discussed for the Model 1.

5.8 Two-Way Mixed Model

Incorporating the sample specific effects as assignable source of variation in the basic assessors' model, we may write a two-way no-interaction additive model as

$$y_{ij} = \mu + \alpha_i + \lambda_j + u_{ij} \,, \tag{5.36}$$

where λ_j represents the deviation from the average score for j^{th} sample. This model would be useful to study the Data Set 6, where the Tasters' scores are given specific to some CTC clones. The clonal effects may be assumed to be fixed and under the assumption of random Tasters' effects, the model (5.36) is typically a mixed effects model. The interaction effect can not be considered here as the Tasters' are independently evaluating single sample from each individual CTC clone. The inclusion of interaction term would have been logical had repeated observations on each clone by each Taster been available.

Note that the model (5.36) is an over-simplified one, which is specific to particular sensory attribute and does not include the possible repetition. The model is written specific to the data under consideration. For n samples and r Tasters, the model may be

written in matrix form as follows :

$$\begin{pmatrix} \underset{\sim}{y}_1 \\ \underset{\sim}{y}_2 \\ \dots \\ \dots \\ \underset{\sim}{y}_r \end{pmatrix} = \mu(1_r \otimes 1_n) + \begin{bmatrix} 1_n & 0 & \dots & \dots & 0 \\ 0 & 1_n & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & 1_n \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \dots \\ \alpha_r \end{bmatrix} + \begin{bmatrix} I_n \\ I_n \\ \dots \\ \dots \\ I_n \end{bmatrix} \begin{pmatrix} \beta_1 \\ \lambda_2 \\ \dots \\ \dots \\ \lambda_n \end{pmatrix} + \begin{pmatrix} \underset{\sim}{u}_1 \\ \underset{\sim}{u}_2 \\ \dots \\ \dots \\ \underset{\sim}{u}_r \end{pmatrix}$$

$$\Rightarrow y = \mu(1_{nr} \otimes 1_n) + Z_\alpha \alpha + Z_\lambda \lambda + u \,, \tag{5.37}$$

where $Z_\alpha = I_r \otimes 1_n$ and $Z_\lambda = 1_r \otimes I_n$ are the matrices of individual dummies associated with Taster specific and sample specific effects respectively. We note here that $Z_\alpha Z_\alpha' = I_r \otimes J_n$ and $Z_\lambda Z_\lambda' = J_r \otimes I_n$.

We assume that $\alpha_i \sim iid(0,\ \sigma_\alpha^2)$, $u_{ij} \sim iid(0,\ \sigma_i^2)$ and λ_j are fixed with $\sum_j \lambda_j = 0$. Also, α_i *and* u_{ij} are independent for all i and j. The assumption on u_{ij} implies that the variation of Tasters in their scoring over samples is unequal. In the model the number of parameters to be estimated is (n+r+2), as there are n fixed effects, one location parameter and (r+1) variance components.

In the following two sections we discuss the ANOVA and ML estimators of parameters in the model (5.36).

5.8.1 ANOVA Estimation

For the two-way model (5.36), we have

$$\begin{aligned} y_{i0} &= \frac{1}{n}\sum_j y_{ij} = \mu + \alpha_i + u_{i0} \\ y_{ij} &= \frac{1}{r}\sum_i y_{ij} = \mu + \alpha_0 + \lambda_j + u_{0j} \\ y_{00} &= \frac{1}{n}\sum_j y_{0j} = \mu + \alpha_0 + u_{00},\ \ \alpha_0 = \frac{1}{n}\sum_i \alpha_i \,. \end{aligned}$$

$$\begin{aligned} E(SSe) &= \sum_i \sum_j E\left[(u_{ij} - u_{i0}) - (u_{0j} - u_{00})\right]^2 \\ E(u_{ij} - u_{i0})^2 &= E(u_{ij}^2) + E(\frac{1}{n}\sum_j u_{ij})^2 - 2E(u_{ij}\frac{1}{n}\sum_j u_{ij}) \end{aligned}$$

$$= \sigma_i^2 + \frac{1}{n^2} \sum_j E(u_{ij}^2) - \frac{2}{n} E(u_{ij}^2) = (1 - \frac{1}{n}) \sigma_i^2$$

$$E(u_{0j} - u_{00})^2 = E(\frac{1}{r} \sum_i u_{ij})^2 + \frac{1}{r^2} \sum_i E(\frac{1}{n} \sum_j u_{ij})^2 - 2E(\frac{1}{r} \sum_i u_{ij} \frac{1}{nr} \sum_i \sum_j u_{ij})$$

$$= \frac{1}{r^2} \sum_i \sigma_i^2 + \frac{1}{n^2 r^2} \sum_i \sum_j \sigma_i^2 - \frac{2}{n r^2} \sum_i \sigma_i^2$$

$$= \frac{1}{r^2} (1 - \frac{1}{n}) \sum_i \sigma_i^2$$

$$E(u_{ij} - u_{i0})(u_{0j} - u_{00}) = E(u_{ij} \frac{1}{r} \sum_i u_{ij}) - E(u_{ij} \frac{1}{nr} \sum_i \sum_j u_{ij})$$

$$-E(\frac{1}{n} \sum_j u_{ij} \frac{1}{r} \sum_i u_{ij}) + E(\frac{1}{n} \sum_j u_{ij} \frac{1}{nr} \sum_i \sum_j u_{ij})$$

$$= \frac{1}{r} \sigma_i^2 - \frac{1}{n r} \sigma_i^2 - \frac{1}{n r} \sigma_i^2 + \frac{1}{n r} \sigma_i^2 = \frac{1}{r} (1 - \frac{1}{n}) \sigma_i^2 .$$

Thus,

$$\begin{aligned} E(SSe) &= \textstyle\sum_i \sum_j \left[(1 - \frac{1}{n}) \sigma_i^2 + \frac{1}{r^2} (1 - \frac{1}{n}) \sum_i \sigma_i^2 - 2 \frac{1}{r} (1 - \frac{1}{n}) \sigma_i^2\right] \\ &= (n-1)(1 - \tfrac{1}{r}) \textstyle\sum_i \sigma_i^2 \end{aligned} \tag{5.38}$$

$$E(SS\alpha) = \sum_i \sum_j \left[E(\alpha_i - \alpha_0)^2 + E(u_{i0} - u_{00})^2\right]$$

$$E(\alpha_i - \alpha_0)^2 = E(\alpha_i^2) + E(\frac{1}{r} \sum_i \alpha_i)^2 - 2E(\alpha_i \frac{1}{r} \sum_i \alpha_i)$$

$$= (1 - \frac{1}{r}) \sigma_\alpha^2$$

$$E(u_{i0} - u_{00})^2 = E(\frac{1}{n} \sum_j u_{ij})^2 + E(\frac{1}{n r} \sum_i \sum_j u_{ij})^2 - 2E(\frac{1}{n} \sum u_{ij} \frac{1}{n r} \sum_i \sum_j u_{ij})$$

$$= \frac{1}{n} \sigma_i^2 + \frac{1}{n^2 r^2} \sum_i n \sigma_i^2 - \frac{2}{n r} \sigma_i^2$$

$$= \frac{1}{n} (1 - \frac{2}{r}) \sigma_i^2 + \frac{1}{n r^2} \sum_i \sigma_i^2$$

Thus,

$$E(SS\alpha) = n(r-1) \sigma_\alpha^2 + (1 - \frac{1}{r}) \sum_i \sigma_i^2 , \tag{5.39}$$

From (5.39), equating SSe to $E(SSe)$ we get

$$\frac{1}{n-1} SSe = (1 - \frac{1}{r}) \sum_i \sigma_i^2$$

$$\Rightarrow \sum_i \sigma_i^2 = \frac{1}{n-1}\,\frac{r}{r-1}\,SSe \tag{5.40}$$

From (5.39), equating $SS\alpha$ to $E(SS\alpha)$ and using (5.40) we get

$$SS\alpha = n(r-1)\,\sigma_\alpha^2 + \frac{1}{n-1}\,SSe$$

$$\Rightarrow \sigma_\alpha^2 = \frac{1}{n\,(r-1)}\left[SS\alpha - \frac{1}{n-1}\,SSe\right] \tag{5.42}$$

Again,

$$\begin{aligned} E(SS\lambda) &= SS\lambda \\ \Rightarrow E\left[\sum_i \sum_j (\lambda_j + u_{0j} - u_{00})\right]^2 &= SS\lambda \\ \Rightarrow \sum_i \sum_j \lambda_j^2 + \sum_i \sum_j E(u_{0j} - u_{00})^2 + 0 &= SS\lambda \end{aligned}$$

$$\begin{aligned} E(u_{oj} - u_{00})^2 &= E(u_{0j}^2) + E(u_{00}^2) - 2E(u_{0j}\,u_{00}) \\ &= \frac{1}{r^2}\,(1 - \frac{1}{n}) \sum_i \sigma_i^2 \end{aligned}$$

Thus,

$$\begin{aligned} \sum_i \sum_j \lambda_j^2 + \frac{1}{r}\,(n-1) \sum_i \sigma_i^2 &= SS\lambda \\ \Rightarrow r \sum_j \lambda_j^2 &= SS\lambda - \frac{1}{r}\,(n-1)\,\frac{1}{n-1}\,\frac{r}{r-1}\,SSe \\ \Rightarrow \sum_j \lambda_j^2 &= \frac{1}{r}\,SS\lambda - \frac{1}{r(r-1)}\,SSe \end{aligned}$$

Now for fixed i,

$$\begin{aligned} \sum_j (y_{ij} - y_{i0})^2 &= \sum_j [\lambda_j + (u_{ij} - u_{i0})]^2 \\ E\left[\sum_j (y_{ij} - y_{i0})^2\right] &= \sum_j \lambda_j^2 + \sum_j E\,(u_{ij} - u_{i0})^2 \\ &= \frac{1}{r}\,SS\lambda - \frac{1}{r(r-1)}\,SSe + (n-1)\,\sigma_i^2 \end{aligned}$$

Thus, σ_i^2 may be obtained from the following relation

$$(n-1)\,\sigma_i^2 + \frac{1}{r}\,SS\lambda - \frac{1}{r(r-1)}\,SSe = \sum_j (y_{ij} - y_{i0})^2 \tag{5.42}$$

5.8.2 ML Estimation

Under the assumption of normality of response, the likelihood function for the two-way mixed model may be written as

$$L = (2\Pi)^{-\frac{nr}{2}} \left(\prod_i \sigma_i^2\right)^{-\frac{n-1}{2}} \left[\prod_i (\sigma_i^2 + n\sigma_\alpha^2)\right]^{-\frac{1}{2}}$$

$$exp\left[-\frac{1}{2}(y - \mu\, 1_{nr} - Z_\lambda\, \lambda)'\, \Phi^{-1}\, (y - \mu\, 1_{nr} - Z_\lambda\, \lambda)\right], \qquad (5.43)$$

and the quadratic form in the exponent may be written as

$$\sum_i (y_i - \mu\, 1_n - I_n\, \lambda)' \left[\frac{1}{\sigma_i^2} I_n - \frac{1}{\sigma_i^2} \frac{\sigma_\alpha^2}{\sigma_i^2 + n\sigma_\alpha^2} J_n\right] (y_i - \mu\, 1_n - I_n\, \lambda)$$

$$= \sum_i \frac{1}{\sigma_i^2} \sum_j (y_{ij} - \mu - \lambda_j)^2 - \sum_i \frac{1}{\sigma_i^2} \frac{\sigma_\alpha^2}{\sigma_i^2 + n\sigma_\alpha^2} \left[\sum_j (y_{ij} - \mu - \lambda_j)\right]^2 .$$

Note that λ_j being fixed, the dispersion matrix Φ has the same form as that of Φ_1 defined in (5.15).

The loglikelihood function may be written as

$$l = -\tfrac{nr}{2} ln\,(2\,\pi) - \tfrac{n-1}{2} \textstyle\sum_i ln\, \sigma_i^2 - \tfrac{1}{2} \sum_i ln\,(\sigma_i^2 + n\,\sigma_\alpha^2) - \tfrac{1}{2} \sum_i \sum_j \tfrac{1}{\sigma_i^2} (y_{ij} - \mu - \lambda_i)^2$$
$$+\tfrac{1}{2} \textstyle\sum_i \tfrac{1}{\sigma_i^2} \tfrac{\sigma_\alpha^2}{\sigma_i^2 + n\sigma_\alpha^2} \left[\sum_j (y_{ij} - \mu - \lambda_j)\right]^2 . \qquad (5.44)$$

Differentiating (5.44) with respect to μ and equating to zero, the ML estimator of location parameter is obtained as

$$\mu = \frac{1}{n} \sum_j y_j^* \quad \text{where} \quad y_j^* = \frac{\sum_i w_i\, y_{ij}}{\sum w_i}, \quad w_i = \frac{1}{\sigma_i^2 + n\sigma_\alpha^2} .$$

The estimator of fixed parameter λ_j may be obtained from the following relation

$$\begin{aligned} \tfrac{\partial l}{\partial \lambda_j} = 0 \;&\Rightarrow\; \textstyle\sum_i \tfrac{1}{\sigma_i^2} (y_{ij} - \mu - \lambda_j) - \sum_i \tfrac{1}{\sigma_i^2} \tfrac{\sigma_\alpha^2}{n\sigma_i^2 + \sigma_\alpha^2} \sum_j (y_{ij} - \mu - \lambda_j) = 0 \\ &\Rightarrow\; \textstyle\sum_i \tfrac{1}{\sigma_i^2} \lambda_j = \sum_i \tfrac{1}{\sigma_i^2} (y_{ij} - \mu) - \sum_i \tfrac{1}{\sigma_i^2} \tfrac{\sigma_\alpha^2}{n\sigma_i^2 + \sigma_\alpha^2} \sum_j (y_{ij} - \mu) \qquad (5.45) \\ &\Rightarrow\; \lambda_j = (\bar{y}_{iw} - \mu) - (\textstyle\sum_i \tfrac{1}{\sigma_i^2})^{-1} \sum_i \tfrac{\sigma_\alpha^2}{\sigma_i^2(\sigma_i^2 + n\sigma_\alpha^2)} \sum_j (y_{ij} - \mu) , \end{aligned}$$

where $\bar{y}_{iw} = \frac{\sum_i a_i\, y_{ij}}{\sum_j a_i}$, $a_i = \frac{1}{\sigma_i^2}$. Clearly, $\bar{y}_{iw}$ is the weighted mean where weights are reciprocal of the error variances associated with individual Tasters.

Also, the estimators of variances components σ_α^2 and σ_i^2 may be obtained from the following relations respectively,

$$\sum_i \frac{1}{\sigma_i^2 + n\sigma_\alpha^2} \left[1 - \frac{n}{\sigma_i^2 + n\sigma_\alpha^2} A_i^2\right] = 0 \tag{5.46}$$

$$\frac{\sigma_i^2}{\sigma_i^2 + n\sigma_\alpha^2} + \frac{\sigma_\alpha^2}{\sigma_i^2 + n\sigma_\alpha^2} \left[\frac{1}{\sigma_i^2} + \frac{1}{\sigma_i^2 + n\sigma_\alpha^2}\right] A_i^2 = \frac{1}{\sigma_i^2} \sum_j (y_{ij} - \lambda_j - \mu)^2 - (n-1), \tag{5.47}$$

where $A_i = \sum_i (y_{ij} - \mu - \lambda_j) = \sum_j (y_{ij} - \mu)$.

5.9 Data Analysis

5.9.1 Analysis of Data Set 5

A set of 14 CTC samples are evaluated independently by a panel of three experienced Tasters in terms of 'strength', 'quality' and 'overall quality'. All the Tasters used the same structured scale and evaluated the samples on 0-10 point scale. The basic statistics for the three different attributes are presented in Chapter 1. We note here that the score on each sample is the average of 10 repeats. However the repeated scores are not provided to us.

As evident from the basic statistics, the scores are minimum (on average) on strength and are highest for the overall quality or value. We first perform the two-way ANOVA without interaction on the three attributes separately. For all these attributes, the 'between-Taster' variation has come out highly significant at 5% level. However, the within-sample variability is insignificant. The ANOVA results are presented in Table 5.1.

We have tested the significance of difference among the average scores under the assumption of unequal error variance. The average score of the three Tasters differ significantly for S and V. However, for Q the difference is insignificant. The profile plot of the scores on S, Q and V are presented in Figure 5.1 to Figure 5.3.

Since the within sample variability is insignificant for all the three attributes, we may use the Model 1 and/or Model 2 to estimate the mean scores and the error variances associated with the three Tasters' scores. We first discuss the ML estimators. We note here that though both Model 1 and 2 are tried for the given sets of data, we only present

the estimates obtained using Model 1, as this model provides better fit, on the basis of likelihood information and the sum of squares of errors. The ML estimates of location parameter and the variance components for all the three attributes are presented in Table 5.2 and Table 5.4 contains the estimated scores for the attributes along with the sum of squared distance between the estimated scores and the three Tasters' scores.

The estimates of σ_1^2, σ_2^2 and σ_3^2 may be considered to be the guiding factor to decide upon the precision of individual Taster's. However, this is a naive approach to assess 'how good' a Taster is. This would favour a Taster scoring consistently within a narrow interval on the scale he/she adopts, whether the Taster is able to separate or distinguish the samples or not. Here σ_i^2s concerns precision actually. If the samples are truly random having the same intrinsic properties, then σ_i^2 would measure how inconsistent the i^{th} Taster is in evaluating the same type of samples. For moderately large sample size we could use the Burtlett's test (Judge et al., 1995, p. 448) for equality of σ_i^2s or we could easily develop a likelihood ratio test to test the hypothesis : $H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_r^2$ and $H_0 : \sigma_\alpha^2 = 0$, under H_0. We do not opt for either of these two tests because of very small sample size ($n = 14$). However, from Table 5.2 it appears that σ_i^2s are different, especially for 'strength' and 'value'.

As may be observed from Table 5.2 the patterns in the estimates of σ_i^2s for three different attributes is not the same, though the same Tasters have evaluated the samples. For V the Taster 3 has the highest error variance whereas Taster 1 is having the least. But for Q the error variance is maximum for Taster 1 and is least for the second Taster. A completely different pattern is observed for S.

The estimates of σ_α^2s for the three attributes are relatively small, suggesting that the three Tasters do agree on average for all the three attributes. In fact, the estimates of σ_α^2 for 'strength' and 'value' are close to zero. This, in turn, support our approach of proposing the Model 1 for the given sets of data.

In Table 5.4, $\sum_{j=1}^{14}(\hat{y}_j - y_{ij})^2$ represents the sum of squared distance between the estimated scores and the i^{th} Taster's sensory scores. Higher the value of $\sigma_i^2 (i = 1, 2, 3)$, higher would be the value for $\sum_j (\hat{y}_{ij} - y_{ij})^2$. A typical feature of the scores on V is that the estimated scores for last 7 samples are all lower than these for the first of samples. But such feature

is not observed for the other two attributes.

5.9.2 Analysis of Data Set 6

A set of 16 CTC samples are evaluated by a panel of five Tasters independently in terms of 'strength' and 'quality'. These samples are developed through blending of different CTC clones in different proportions and thus the quality characteristic is likely to vary over samples. As mentioned in the introductory chapter, the identification of the clonal combinations are not disclosed.

As may be observed from the two-way analysis of variance result presented in Table 5.5, the variations due to samples as well as Tasters are significant at 5% level. The Tasters' variations for the two attributes are very highly significant. For obvious reasons, we use the two-way heteroscedastic VC model to estimate the error variances associated with the five Tasters' scores along with the mean scores. The sample specific effects may also be estimated.

The ML and ANOVA estimators of variance components and the mean scores are presented in Table 5.6 and Table 5.7 respectively, along with the estimated loglikelihood values for the two attributes. It may be observed that the ML estimates of error variances associated with the Tasters (σ_i^2 s) are not very large. However, the estimates of σ_α^2 are comparatively large, especially for the attribute quality. This implies disagreement among the Tasters in their average choices. Small values of σ_i^2 implies that the individual Tasters' choices do not vary much over the different clonal combinations. We note here that the estimates of σ_i^2 can not be considered as the guiding factor to decide upon the precision of the Tasters in this case. This is because each sample represents a particular clonal combination and the variation in terms of quality attribute(s) over samples is most likely. Small value of $\hat{\sigma_i^2}$ implies that the i^{th} Taster do not find much difference in the samples in terms of strength and quality. No specific interpretation about the samples can be drawn on the basis of these findings, as only a single sample for each clonal combination is studied. Had several observations on each clonal combination been available, we could infer about the characteristics of the samples with validity. For the same reason of poor information, we can not introduce much discussion on the estimates of fixed effects parameters λ_j.

The estimated scores $(\hat{\mu} + \hat{\lambda_j})$ and the estimates of λ_j along the row scores of the five Tasters on strength and quality are presented in Table 5.8 and Table 5.9. It may be observed from the row scores that the scores given by the fourth Taster (T 4) on quality are very low as compared to those given by other. In fact this Taster's scores ranges from 3.71 to 4.50, which is far below the ranges of the other Tasters' scores. Profile plots of Tasters' choices on strength and quality along with the estimated scores are presented in Figure 5.4 and 5.5.

Table 5.1 : Two-Way ANOVA result

	Source	*df*	*SS*	*F*	F critical at 5%
Strength	Sample	13	4.38	1.38	2.12
	Taster	2	6.24	12.83	3.37
	Error	26	6.32		
Quality	Sample	13	3.22		2.12
	Taster	2	13.52		3.37
	Error	26	4.74		
Value	Sample	13	11.28	2.33	2.12
	Taster	2	6.55	8.78	3.37
	Error	26	9.70		

Table 5.2 : ML estimators of μ and variance components

	Strength	Quality	value
$\hat{\mu}$	5.2299	6.3909	7.8867
$\hat{\sigma}_1^2$	0.5772	0.2217	0.0684
$\hat{\sigma}_2^2$	1.1980	0.1154	0.6157
$\hat{\sigma}_3^2$	0.0001	0.1658	1.4104
$\hat{\sigma}_\alpha^2$	0.5772	0.2217	0.0684
$ln\ L$	-19.4351	-19.0047	-21.5630

Table 5.3 : ANOVA estimators of μ and variance components

	Strength	Quality	value
$\hat{\mu}$	5.4302	6.4987	7.7854
$\hat{\sigma}_1^2$	0.6925	0.2943	0.1358
$\hat{\sigma}_2^2$	1.3645	0.2541	0.6352
$\hat{\sigma}_3^2$	0.0021	0.1956	1.5684
$\hat{\sigma}_\alpha^2$	0.9824	0.3212	0.1653

Table 5.4 : Estimated scores for the three attributes

Sample	Strength	Quality	value
1	5.75	6.94	8.21
2	4.84	6.68	8.08
3	5.04	7.09	8.28
4	4.25	6.92	8.10
5	4.15	6.73	8.09
6	4.50	6.76	8.27
7	4.85	6.67	8.12
8	4.00	5.85	7.73
9	4.15	6.07	7.74
10	4.28	6.19	7.60
11	4.35	5.81	7.35
12	3.85	5.88	7.69
13	4.25	6.09	7.57
14	4.55	5.69	7.58
$\sum_{(} \hat{y}_j - y_{1j})^2$	8.07	2.49	0.68
$\sum(\hat{y}_j - y_{2j})^2$	16.77	1.01	8.35
$\sum(\hat{y}_j - y_{3j})^2$	0.00001	1.71	19.47

Table 5.5 : Two-Way ANOVA for Data Set 6

	Source	*df*	*SS*	*F*	F critical at 5%
Strength	Sample	15	3.68	0.68	1.84
	Taster	4	44.61	30.87	2.53
	Error	60	21.68		
Quality	Sample	15	4.09	1.35	1.84
	Taster	4	136.32	169.69	2.53
	Error	60	12.08		

Table 5.6 : ML estimators of μ and variance components

Parameters	Strength	Quality
$\hat{\mu}$	6.7433	6.3379
$\hat{\sigma}_1^2$	0.1173	0.1285
$\hat{\sigma}_2^2$	0.5957	0.4888
$\hat{\sigma}_3^2$	0.0783	0.2003
$\hat{\sigma}_4^2$	0.8167	0.1261
$\hat{\sigma}_5^2$	0.1966	0.2055
$\hat{\sigma}_\alpha^2$	0.5029	1.6858
$\hat{ln\, L}$	-32.0607	-30.5576

Table 5.7 : ANOVA estimators

	Strength	*Quality*
$\hat{\mu}$	6.9985	6.7568
$\hat{\sigma}_1^2$	0.2365	0.3212
$\hat{\sigma}_2^2$	0.6984	0.5878
$\hat{\sigma}_3^2$	0.0985	0.3012
$\hat{\sigma}_4^2$	0.9987	0.1546
$\hat{\sigma}_5^2$	0.2122	0.3214
$\hat{\sigma}_\alpha^2$	0.7561	1.8751

Table 5.8 : The Tasters' scores on strength and estimated scores along with the fixed effects estimates $\hat{\lambda_j}$

T 1	T 2	T 3	T 4	T 5	Estimated Score	λ_j
8.00	4.85	7.43	6.57	8.14	7.1815	0.4382
7.85	5.28	7.00	7.28	7.14	6.8310	0.08771
7.00	4.85	7.28	6.74	7.43	6.6975	-0.0458
7.75	4.86	7.14	6.14	7.85	6.8621	0.11884
7.29	5.14	7.29	6.57	7.43	6.8071	0.06389
7.14	5.14	6.86	6.71	7.00	6.5071	-0.2362
7.33	5.00	7.00	7.00	7.17	6.6571	-0.0862
7.17	4.50	7.17	6.67	7.67	6.7290	-0.0143
7.50	5.50	7.25	7.00	7.75	6.9449	0.2017
7.50	6.00	6.50	3.50	7.25	6.4134	-0.3299
7.25	6.50	7.25	6.50	7.25	6.8218	0.07850
7.75	7.25	7.50	5.50	7.00	7.0341	0.29081
7.00	5.25	7.25	6.50	7.50	6.7208	-0.0225
7.25	4.75	6.75	7.00	8.00	6.6547	-0.0887
7.50	5.57	7.00	6.25	6.50	6.6018	-0.1415
6.75	6.25	6.75	6.00	7.00	6.3795	-0.3638

Table 5.9 : The Tasters' scores on quality and estimated scores along with the fixed effects estimates $\hat{\lambda_j}$

T 1	T 2	T 3	T 4	T 5	Estimated Score	λ_j
7.71	5.00	7.71	4.14	8.00	7.0125	0.6746
7.57	5.14	7.00	4.43	7.28	6.9062	0.5683
7.00	4.85	7.71	4.28	7.57	6.7986	0.4607
7.57	4.57	7.71	4.14	7.85	6.9496	0.6117
7.29	4.86	7.71	4.40	7.43	6.9585	0.6206
7.14	4.71	6.86	3.71	7.00	6.4798	0.1419
7.33	4.67	7.50	3.83	7.50	6.7219	0.3814
7.17	7.17	7.50	3.83	7.67	6.6289	0.2909
7.50	5.50	7.50	4.25	7.50	6.9659	0.6280
7.50	6.00	6.75	4.50	7.00	6.9214	0.5835
7.25	6.50	8.00	4.50	7.25	7.1563	0.8184
8.25	7.00	7.50	4.50	7.00	7.4211	1.0832
7.00	5.25	7.00	4.25	8.00	6.6822	0.3443
7.25	4.75	7.00	5.00	8.00	6.9775	0.6395
7.50	5.50	6.50	4.00	6.50	6.6407	0.3028
6.75	6.00	7.00	3.75	7.00	6.4687	0.1308

Figure 5.1: Profile plot of scores on strength

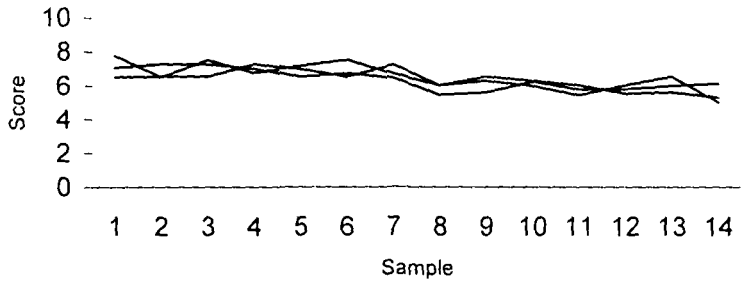


Figure 5.2: Profile plot of scores on quality

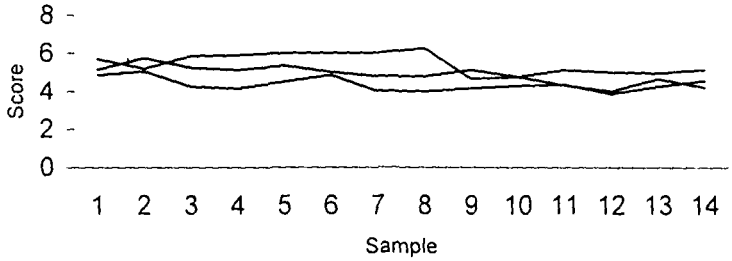
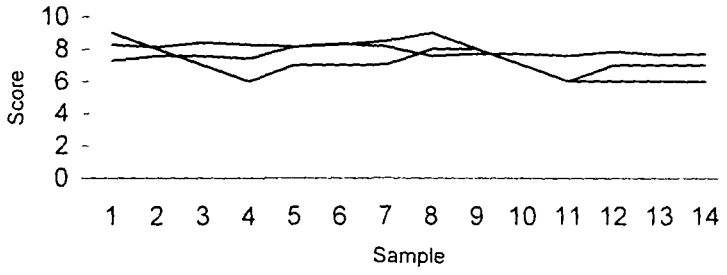


Figure 5.3: Profile plot of scores on value



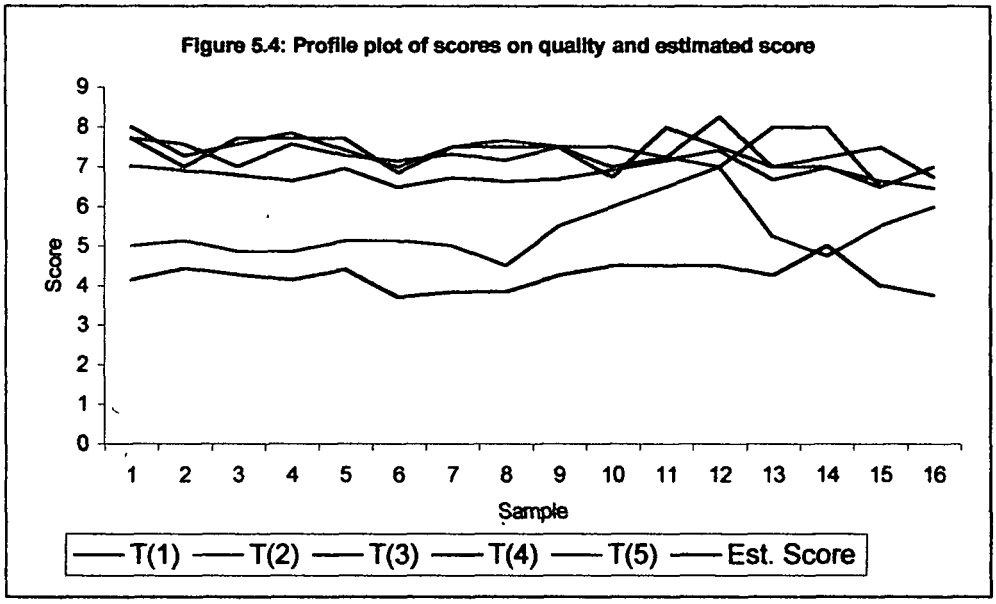


Figure 5.4: Profile plot of scores on quality and estimated score
Score
Sample
T(1) T(2) T(3) T(4) T(5) Est. Score

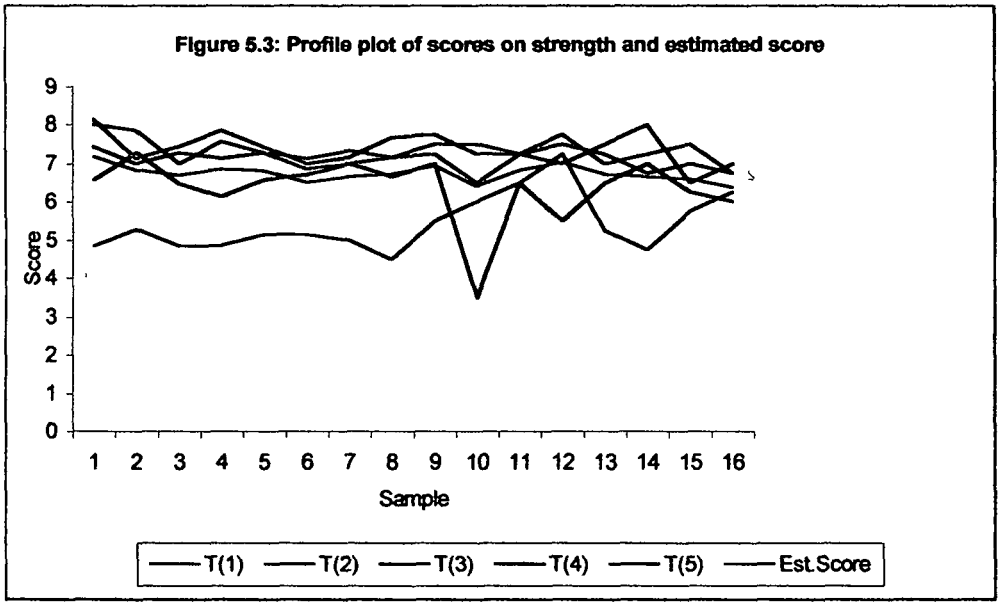


Figure 5.3: Profile plot of scores on strength and estimated score
Score
Sample
T(1) T(2) T(3) T(4) T(5) Est.Score

CHAPTER - 6

ONE-WAY ERROR COMPONENT REGRESSION MODELS WITH HETEROSCEDASTIC ERROR

6.1 Introduction

In the third chapter we discussed the association of biochemical quality parameters in tea with the sensory score given by single Taster. In this chapter we study the possibility of associating the chemical information in tea samples with sensory scores given by a panel of Tasters. We adopt the regression approach here. In the regression setup the measured values of chemical parameters are treated as regressors and these measurements are known to be obtained with high degree of accuracy. The dependent variable (response) is nothing but the sensory scores given by the panel of Tasters independently. Thus we have repeated observations on the response variate.

The problem may be viewed typically as a chemometric one, where the information set includes both chemical parameters and the sensory scores. Note that the measurements on biochemical parameters are fixed for a particular sample. Only the sensory choices vary due to Tasters. Unlike a typical sensory panel data, we have no replicated scores by Tasters on each sample. Only a single score on each sample (for a particular attribute) by a particular Taster is available. Thus we can not study the different possible aspects of the sensory panel data related to 'scale differences', 'non-linearity component' for sample $\times$ assessor 'interaction' etc., while relating chemical information with the sensory scores. Given the Data Sets 7 and 8, we can atb est study the individual variations among the Tasters apart from identifying the biochemical quality parameters, which are statistically significant in terms of influencing a particular quality attribute.

One possible statistical approach in this situation may be the multiple linear regression on the means of the scores given by the Tasters. In this approach we ignore the subjectivity associated with the assessors, choices and can not track the individual variation. Again, if the regressors are interrelated leading to multicollinearity problem, the applications of techniques like Principal Component Regressions may be used. However, for the data sets under study, the conditional Index test does not indicate severe multicollinearity.

Keeping in view the inherent subjectivity associated with the Tasters' choices, we may introduce an error component with the response variable. As several Tasters' scores are available for a given set of samples, we may formulat our regression model so as to take into account the individual variations due to Tasters. Keeping this in view we propose some error-component regression models in this chapter. As has been discussed in the last chapter, there is a continuing debate on whether to assume Tasters' affects as fixed or random. In the error component regression set up, we discuss random effects formulations, under the assumption that the error variances associated with different Tasters' scores are unequal. Thus our study is specific to error-component regression models with heteroscedastic errors.

The error-component models are well developed in the statistical literature. (Baltagi, 1996) and we do not claim any originality in our study, so for the statistical model formulation is concerned. We simply extend the basic model formulation and discuss the estimation procedures in a heteroscedastic situation.

Before we formally introduce the error-component models, we need to undertake some basic diagnostics, which is discussed in the next section.

6.2 The Basic Diagnostics

In the regression set up, under the assumption of same distribution for different individual responses, it is important to study whether the intercepts and/or the regression coefficients exhibit same pattern while associating the chemical information on a particular set of tea samples with different Tasters' scores. Here we address the problem of studying

the stability (consistency) of different coefficient estimates. If, for the given two data sets under study, it is established that the sets of regression coefficients are different for different individual responses, then we may have to incorporate this into the model in an appropriate manner.

In a multivariate multiple regression setup, we model the relationship between r responses $\underset{\sim}{y}_1, \underset{\sim}{y}_2, \ldots, \underset{\sim}{y}_r$, and a single set of predictor variables, $x_1, x_2, \ldots x_k$, as

$$\begin{aligned}
\underset{\sim}{y}_1 &= \mu_1 + \beta_{11}x_1 + \ldots + \beta_{k1}x_k + u_1 \\
\underset{\sim}{y}_2 &= \mu_2 + \beta_{12}x_1 + \ldots + \beta_{k2}x_k + u_2 \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\underset{\sim}{y}_r &= \mu_r + \beta_{1r}x_1 + \ldots + \beta_{kr}x_k + u_r \,. \qquad (6.1)
\end{aligned}$$

Each response is assumed to follow its own regression model. The error component $u = (u_1, u_2, \ldots u_r)'$ has $E(u) = 0$ and $Var(u) = \Phi$. The error component associated with different responses may be correlated. In matrix form the above formulation may be written as $y = X\beta + u$, where

$$Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1r} \\ y_{21} & y_{22} & \cdots & y_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ y_{n1} & y_{n2} & \cdots & y_{nr} \end{bmatrix}_{n\times r}, \quad \beta = \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_r \\ \beta_{11} & \beta_{12} & \cdots & \beta_{1r} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \beta_{k1} & \beta_{k2} & \cdots & \beta_{kr} \end{bmatrix}_{(k+1)\times r}$$

$$u = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1r} \\ u_{21} & u_{22} & \cdots & u_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ u_{n1} & u_{n2} & \cdots & u_{nr} \end{bmatrix}$$

Here $E(\underset{\sim}{u}_i) = 0$, $Cov\,(\underset{\sim}{u}_i, \underset{\sim}{u}_l) = \sigma_{ie}\, I$ for $i, l = 1, 2, \ldots r$. We are primarily interested in the estimates of the regression coefficients vectors $\underset{\sim}{\beta}_i$. Note that for the i^{th} response, the

GLS estimator of $\underset{\sim}{\beta}_i$ would be $(X'\ \Sigma_i^{-1}\ X)^{-1}\ X'\underset{\sim}{y}_i$, where $\Sigma_i = \sigma_{ii}\ I$. Thus for all the responses, the estimated GLS estimator would be

$$\hat{\beta} = \left[\underset{\sim}{\hat{\beta}}_1 \underset{\sim}{\hat{\beta}}_2 \ldots \underset{\sim}{\hat{\beta}}_r\right] = (X'\ \hat{\Phi}^{-1}\ X)^{-1}\ X'\hat{\Phi}^{-1}\left[\underset{\sim}{y}_1 \underset{\sim}{y}_2 \ldots \underset{\sim}{y}_r\right],$$

which is obtained by minimizing $u_i'\ \Phi^{-1}\ u_i = (y_i - X\underset{\sim}{\beta}_i)'\ \Phi^{-1}(y_i - X\underset{\sim}{\beta}_i)$ for each i. Here $\hat{\beta}$ is unbiased and consistent with covariance matrix $E[(\underset{\sim}{\hat{\beta}} - \underset{\sim}{\beta})(\underset{\sim}{\hat{\beta}} - \underset{\sim}{\beta})'] = [X'\ \Phi^{-1}\ X]^{-1}$. To obtain the estimator $\hat{\sum}$ we use the OLS residuals $\hat{u}_i = \underset{\sim}{y}_i - X_i\ \hat{\beta}_i$. The i^{th} element in $\hat{\sum}$ is of the form $\hat{\sigma}_{ie} = \frac{1}{n}\ \hat{u}_i'\ \hat{u}_l$.

Before we develop a formal error-component regression model combining all the responses together, it is important to test if $\underset{\sim}{\beta}_1, \underset{\sim}{\beta}_2, \ldots \underset{\sim}{\beta}_r$, the vectors of regression coefficients are all equal for the r responses on the given set of n samples. Here the problem is to study the stability of regression coefficients for different responses given the same set of predictors for all the responses. This can be done introducing dummy variables in the regression model. The techniques and implications of introducing dummy variables to allow for differences in the intercepts and/or slopes are well developed in the econometric literature. To test the hypothesis of stability (or consistency) of β coefficients, we may extend the ANOVA test proposed by Chow (1960). In our situation, the null hypothesis to be tested is

$$H_O : \beta_{11} = \ldots = \beta_{1r}, \ldots,\ \ \beta_{k1} = \ldots = \beta_{kr} \text{ and } \mu_1 = \mu_2 = \ldots = \mu_r\ .$$

If this hypothesis holds, we can estimate a single equation pooling the responses together and assume fixed regression coefficients for all the independent response variates. Otherwise, we have to develop regression model with varying slopes over the responses. Here we use the concept of restricted error sum of squares (RESS) and the unrestricted error sum of squares (UESS). To attain UESS we estimate the regression model for each of the equations defined in (6.1). Suppose UESS_i is the error sum of squares for the i^{th} equation with regression error variance σ^2. Then $\frac{UESS_i}{\sigma^2}$ follows χ^2 distribution with $n_i - k$ degrees of freedom, where n_i is the sample size for i^{th} equation. Now since the responses are

independently made by panel of Tasters, $\sum_{i=1}^{r} URSS_i/\sigma^2$ has a χ^2 distribution with d.f. $\sum_{i=1}^{r} n_i - rk - r$. We have, $URSS = \sum_{i=1}^{r} URSS_i/\sigma^2$.

Again, RESS is obtained from the regression with the pooled data, which obviously imposes the restriction that the regression parameters are the same (H_O). It is known that $RESS/\sigma^2$ has a χ^2 distribution with d.f. $\sum_{i=1}^{r} n_i - k - 1$. Now the ratio

$$F = \frac{(RESS - UESS)/(r-1)(k+1)}{UESS/(\sum_i n_i - rk - r)} \tag{6.2}$$

has a F-distribution with d.f. $(r-1)(k+1)$ and $\sum_{i} n_i - rk - r$. Note that the term $(r-1)(k+1)$ represents the difference of the degrees of freedom of RESS and UESS. One limitation of this F-test is that it gives a general test about the equality of all the slope coefficients and intercepts. This. Chow test might tell the consistency of all the coefficients estimates but not tell us which particular coefficients are inconsistent. Keeping this problem in view, the use of dummy variables in the regression model may be suggestive and with this approach are may check for the significance of different dummy variables looking at the t-ratios. But it the number of responses is large, then there will be too many dummy variables in the model. Also if the multicollinearity problem exists, then the t-ratios for each of the regression coefficients are likely to be insignificant and still the F-ratio for the entire set of coefficients is significant. That is why we propose the use of F-test first.

For both the data sets under study, the estimated values of F do not suggest rejection of the null hypothesis. These estimates we presented in the data analysis section. On the basis of our diagnostics. We consider the error-component regression models, which are described in the section 6.4. A brief discussion on the response variable error is followed in the following section.

independently made by panel of Tasters, $\sum_{i=1}^{r} URSS_i/\sigma^2$ has a χ^2 distribution with d.f. $\sum_{i=1}^{r} n_i - rk - r$. We have, $URSS = \sum_{i=1}^{r} URSS_i/\sigma^2$.

Again, RESS is obtained from the regression with the pooled data, which obviously imposes the restriction that the regression parameters are the same (H_O). It is known that $RESS/\sigma^2$ has a χ^2 distribution with d.f. $\sum_{i=1}^{r} n_i - k - 1$. Now the ratio

$$F = \frac{(RESS - UESS)/(r-1)(k+1)}{UESS/(\sum_i n_i - rk - r)} \tag{6.2}$$

has a F-distribution with d.f. $(r-1)(k+1)$ and $\sum_i n_i - rk - r$. Note that the term $(r-1)(k+1)$ represents the difference of the degrees of freedom of RESS and UESS.

One limitation of this F-test is that it gives a general test about the equality of all the slope coefficients and intercepts. This, Chow test might tell the consistency of all the coefficients estimates but not tell us which particular coefficients are inconsistent. Keeping this problem in view, the use of dummy variables in the regression model may be suggestive and with this approach are may check for the significance of different dummy variables looking at the t-ratios. But it the number of responses is large, then there will be too many dummy variables in the model. Also if the multicollinearity problem exists, then the t-ratios for each of the regression coefficients are likely to be insignificant and still the F-ratio for the entire set of coefficients is significant. That is why we propose the use of F-test first.

For both the data sets under study, the estimated values of F do not suggest rejection of the null h On the basis of our diagnostics. We consider the error-component regression models, which are described in the section 6.4. A brief discussion on the response variable error is followed in the following section.

6.3 Some Discussions on the Response Variable Error

Most of the studies made so far have focused extensive by on problems associated measurement error in independent variate. Carrol et. al. (1995) have discussed dif

examples where either the predictors and/or response variates are measured with errors. Such situations arise in different fields like Nutrition studies, Bioassay studies, etc. Some specific examples are Rudemo (1989) in a Herbiade study, Testensen et al. (1989) in long function in children, heart disease and blood pressure studies by Kannel et al. (1986), Liu and Liang (1992), among many others. Pierce et al. (1992) considered analysis of A-bomb survivor data from the Hiroshima explosion.

Not many studies have been made (in our knowledge) where only the response in measured with error. A clinical example is due to Witter et al.(1989) in which damage of heart muscle cause by a infraction can be assessed accurately, but the procedure is expensive and invasive, and instead it is common practice to use peak cardiac enzyme level in the blood stream as a proxy for the to true response. This is obviously a surrogate response variate.

We may introduce an example from the economic field, where the profit (response) of a company is a function of input price and input quality for a given technology and in such a situation we get measurement error only in response variate. Such situation may arise in Sociological and Psychological studies also.

The extensive attention paid to predictor measurement error is obvious as the predictor measurement error is seldom ignorable. The causes and remedies are studied by many researchers. But the response measurement error is often ignorable, as the model holding for true response hold also for the proxy response, except that a measurement error variance component is added to the response variance. For example, in lines regression models with simple types of response measurement error, the response error is confounded with equation error and the effect is simply to increase the variance of the parameter estimates. Thus response measurement error is ignorable in some cases. However, in most of the empirical situations, the response errors are not ignorable. In more complicated regression models, especially is nonlinear situations, it is important to explicitly account for the responses error in the regression analysis.

Carrol et al. (1993) discuss the unbiased and biased measures of true response. Both additive and multiplicative error structures are considered for a case control study. As

they observe, the case of homoscedastic regression variance provides an example of 'ignorable' response measurement error. In such situations, unless the separate VCs are of independent interest, the response error can be ignored and no repeated or validation data is required. But when the VCs are of independent interest, one must have repeated observation for response variate.

The literature on regression models with repeated observations on response variate is scarce. The cause is possibly the 'no-serious statistical complexity' in the estimation procedure. As discussed earlier, the usual estimators of regression coefficients are still unbiased and consistent. However, if heteroscedasticity is introduced in the repeated observations, it adds some special features to the regression estimators since values of the regressors do not change over repetitions of the regressand. One such result is that the GLS estimates reduces to OLS estimates with the response variate replaced by the weighted mean of repeated observations (Pal and Paul 1997, 1998). One may utilize the repeated observations to estimate the VCs along with improved estimates of regression coefficients.

Problems also arise when we like to test the coefficient estimates for specific values or simply to see the significance of the efficiency of estimates by incorporating distribution assumption on the response error. If both response as well as the equation error are assumed normal, then there is no wayout to isolate these two effects. Identification problem arises for these parameters. Thus the efficiency of the estimates can not be increased. This problem can be partially solved if repeated observations are available for the response measurement.

6.4 One-Way Error Component Regression Model

In this section we consider the one-way error-component linear model, where the Taster specific effects are assumed to be random. The logical explanation behind the assumption of random Tasters' effects has already been given in the previous chapter. Here we present two alternative formations.

The general linear random effects model is of the form $y = X\beta + Z\alpha + e$, which has already been discussed in the second chapter. Here the definition of e is based first on defining $E(y) = X\beta$ and $E(y \mid \alpha) = X\beta + Z\alpha$ and then $e = y - E(y \mid \alpha)$. We generally consider $E(e) = 0$ and $E(\alpha) = 0$. For $Var(\alpha) = A$ and $Var(e) = E$, we have $Var(y) = E + ZAZ'$. The estimators of β and the variance components along with their statistical properties are well developed in the literature (see review chapter).

We formulate our model specific to the tea quality assessment data under study. Suppose for n tea samples, r Tasters have independently evaluated the tea samples. Also, measurements on k biochemical parameters are available and we do not have any missing observation. The tea samples are collected from the same experimental garden, received the same manufacturing/processing system and are assumed to be random samples having the same intrinsic quality. Then we may associate the chemical information with the sensory scores using the following model

$$\begin{aligned} y_{ij} &= \mu + x'_{ij}\,\beta + e_{ij} \\ & \qquad i = 1, 2 \ldots r, \qquad j = 1, 2, \ldots n \end{aligned} \tag{6.3}$$

with i denoting Tasters and j denoting samples, x_{ij} is the matrix of k regressors (biochemical parameters) for the j^{th} sample. Note that x_{ij} is fixed over i. $\beta = (\beta_1, \beta_2, \ldots, \beta_k)'$ is the $k \times 1$ vector of slope coefficients and μ represents the intercept. The error component e_{ij} may be decomposed as

$$e_{ij} = \alpha_i + u_{ij} \tag{6.4}$$

where α_i denotes the unobservable Taster specific effect and u_{ij} denotes the remainder disturbance. Combining (6.3) and (6.4), we have the one-way error component regression model

$$y_{ij} = \mu + x'_{ij}\beta + \alpha_i + u_{ij}\ . \tag{6.5}$$

Here α is the average effect due to i^{th} Taster and u_{ij} varies with Taster and sample. Stacking the data for all the samples, the linear error-component regression model may

be written as

$$\begin{pmatrix} \underset{\sim}{y}_1 \\ \underset{\sim}{y}_2 \\ \cdots \\ \cdots \\ \cdots \\ \underset{\sim}{y}_r \end{pmatrix} = \begin{pmatrix} \mu 1_n \\ \mu 1_n \\ \cdots \\ \cdots \\ \cdots \\ \mu 1_n \end{pmatrix} + \begin{pmatrix} X \\ X \\ \cdots \\ \cdots \\ \cdots \\ X \end{pmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \cdots \\ \cdots \\ \cdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} 1_n & 0 & \ldots & 0 \\ 0 & 1_n & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 1_n \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \cdots \\ \cdots \\ \cdots \\ \alpha_r \end{bmatrix} + \begin{pmatrix} \underset{\sim}{u}_1 \\ \underset{\sim}{u}_2 \\ \cdots \\ \cdots \\ \cdots \\ \underset{\sim}{u}_r \end{pmatrix}$$

$$\Rightarrow y = \mu\,(1_r \otimes 1_n) + \mathbf{X}\beta + Z\alpha + u\,, \tag{6.6}$$

where $\underset{\sim}{y}'_i = (y_{i1}\ y_{i2} \ldots y_{in}), \underset{\sim}{u}'_i = (u_{i1}\ u_{i2} \ldots u_{in})$, $Z = I_r \otimes 1_n$ is the matrix of dummy variables associated with α. X represents the matrix of regressors on n samples for i^{th} repeat and is of the form

$$X = \begin{bmatrix} x_{11} & x_{12} & \ldots & x_{1k} \\ x_{21} & x_{22} & \ldots & x_{2k} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ x_{n1} & x_{n2} & \ldots & x_{nk} \end{bmatrix}$$

Here the X matrix is same for all the repeats. Note that, if α_i are assumed to be fixed Taster specific effects with remainder disturbance stochastic, then we have fixed effects error component model. The x_{ij} are assumed to be independent of u_{ij} for all i and j. The fixed effects (FE) model is an appropriate specification if we are focusing on set of r Tasters. Inference in this case is conditional to the specific r repeats. In FE model, there is a large loss of degrees of freedom as we are estimating r extra fixed effects apart from the location parameters. In fact, we are to include $(r-1)$ dummies in the regression, and too many dummies may aggravate the problem of multicollinearity among the regressors.

Also, it may be mentioned here that the one-way random effects model $y_{ij} = \mu + \alpha_i + u_{ij}$, discussed in the fifth chapter is a special case of the model (6.5) with random α_i when $\beta = 0$ or x_j (the matrix of regressors for j^{th} sample) are all the same.

6.4.1 Heterosedastic Formulation of Dispersion Matrix

Under homoscedasticity, we may introduce the assumptions in RE model : $\alpha_i \sim iid\ (0, \sigma_\alpha^2)$, u_{ij} $iid\ (0, \sigma_u^2)$ and α_i are independent of u_{ij}. In addition, the x_{ij} are independent of α_i and u_{ij} for all i and j. In this situation, the variance-covariance matrix of the composite disturbance $e = Z\alpha + u$ would be of the form

$$\begin{aligned} E(e\,e') &= ZE(\alpha\,\alpha')\,Z' + E(u\,u') \\ &= \sigma_\alpha^2(I_r \otimes J_n) + \sigma_u^2(I_r \otimes I_n) \end{aligned}$$

This implies a homoscedastic variance $Var\ (e_{ij}) = \sigma_\alpha^2 + \sigma_u^2$ for all i and j and an equi-correlated block-diagonal covariance matrix. In fact,

$$\begin{aligned} Cov\ (e_{ij}, e_{lk}) &= \sigma_\alpha^2 + \sigma_u^2\,, \quad \text{for} \quad i = l \ \text{ and } \ j = k \\ &= \sigma_\alpha^2, \quad \text{for} \quad i = l \ \text{ and } \ j \neq k \\ &= 0\,, \quad \text{otherwise} \end{aligned}$$

The best quadratic unbiased estimators of variance components and the ML estimators of the implicit parameters are discussed by Baltagi (1995).

Now the homoscedastic error component model may be generalized to the case where α_i and/or u_{ij} are heteroscedastic. Here we may assume α_i heteroscedastic, i.e., $\alpha_i \sim (0, \sigma_i^2)$ for $i = 1, 2, \ldots r$, but $u_{ij} \sim (0, \sigma_u^2)$. Again we may keep α_i homoscedastic with $\alpha_i \sim (0, \sigma_\alpha^2)$ and impose heteroscedasticity on u_{ij}, i.e., $u_{ij} \sim (0, \sigma_i^2)$. Thus we have two distinct formulations. The one-way regression model with the assumption $\alpha_i \sim iid\ (0, \sigma_\alpha^2)$ and $u_{ij} \sim iid\ (0, \sigma_i^2)$, be termed as *Model I* and that with the assumption $\alpha_i \sim iid\ (0, \sigma_i^2)$ and $u_{ij} \sim iid\ (0, \sigma_e^2)$ be termed as *Model II*.

If we presume that the Tasters' do agree on average about the particular quality attribute for the given set of tea samples, but the variation over samples by each Taster differ, then the use of Model I seems appropriate. Otherwise we may try the Model II. Anyway, the appropriateness of using either of the models may always be tested. Here, Model I and Model II are non-nested and we may use the "information criteria" based test to compare the fit (Vonesh and Chinchilli, 1997).

The heteroscedastic formulation of one-way error component model has already been introduced by Baltagi (1995). Baltagi discussed the estimation of VCs using the OLS residuals $\hat{e}_{ij}$. The estimates proposed are : $\hat{w}_i^2 = \sum_{j=1}^{n} (\hat{e}_{ij} - \hat{e}_{i0}^2/(n-1)$ using the OLS residuals $\hat{e}_{ij}$, and then obtain $\hat{\sigma}_i^2 = \hat{w}_i^2 - \sigma_e^2$. Here the estimate of σ_e^2 is the within residuals MSE, obtained by regression on means. This is clearly a two-stage regression procedure and the OLS regression coefficient estimates are still consistent, but not efficient. This estimation procedure is specific to Model II.

In case of Model I, we have $E(e_{ij}^2) = w_i^2 = \sigma_\alpha^2 + \sigma_i^2$. Using OLS residual $\widehat{e_{ij}}$, we may obtain $\widehat{w_i^2} = \sum_j (\widehat{e_{ij}} - \widehat{e_{i0}})^2/(n-1)$. Also, we may compute $\widehat{\sigma_i^2} = \sum_j (\widehat{e_{ij}} - \widehat{e_{i0}})^2/(n-1)$, using the within residuals. Then $(\widehat{w_i^2} - \widehat{\sigma_i^2})$ gives r estimates of σ_α^2 and $\widehat{\sigma_\alpha^2} = \sum_i (\widehat{w_i^2} - \widehat{\sigma_i^2})/r$ is a consistent estimator of σ_α^2.

Here the problem lies with the fact that the within residual MSE is obtained from the regression on means and may not always provide a stable information if the between Taster variation is very high for a given set of data. This may also effect the standard errors of the estimates of β coefficients. For Model II, this two-stage procedure requires a large n and preferably small r. In our study, we restore to the ML estimation procedure for many obvious reasons. The ML estimates of regression parameters and the VCs are discussed in the following section for both Model I and II.

6.5 ML Estimation For Heteroscedastic Models

6.5.1 ML Estimation For Model I

The variance-covariance matrix of the composite disturbance component $e = Z\alpha + u$, may be presented as

$$\begin{aligned} \Phi_1 = E(e\, e') &= ZE(\alpha\alpha')Z' + E(uu') \\ &= \sigma_\alpha^2 ZZ' + \Sigma_i \otimes I_n = \sigma_\alpha^2 (I_r \otimes J_n) + \Sigma_i \otimes I_n \,, \end{aligned}$$

where $ZZ' = I_r \otimes J_n$ and $\Sigma_i = diag(\sigma_i^2\ \sigma_i^2 \ldots \sigma_i^2)$. We may write, $\Phi_1 = diag(S_1\ S_2 \ldots S_r)$,

where

$$S_i = \begin{bmatrix} \sigma^2+\sigma_i^2 & \sigma^2 & \ldots & \ldots\sigma^2 & \\ \sigma^2 & \sigma^2+\sigma_i^2 & \ldots\ldots & \sigma^2 & \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \sigma^2 & \sigma^2 & \ldots\ldots & \sigma^2+\sigma_i^2 & \end{bmatrix}$$

The determinant and inverse of S_i matrix has already been obtained in Section 5.5.2 of Chapter 5. Under the normality assumption, the probability function of the response y may be written as

$$p(y) = (2\pi)^{-\frac{nr}{2}} \mid \Phi_1 \mid^{-\frac{1}{2}} \; exp[-\frac{1}{2}(y - \mu 1_{nr} - \mathbf{X}\beta)' \, \Phi_1^{-1}(y - \mu 1_{nr} - \mathbf{X}\beta)].$$

The likelihood function for Model I may be written as

$$L_1 = (2\pi)^{-\frac{nr}{2}}(\Pi_i\sigma_i^2)^{\frac{n-1}{2}}[\Pi_i(\sigma_i^2 + n\sigma_\alpha^2)]^{-\frac{1}{2}} exp\,[-\frac{1}{2}(y - \mu 1_{nr} - \mathbf{X}\beta)' \, \Phi_1^{-1}(y - \mu 1_{nr} - \mathbf{X}\beta)],$$

and the quadratic form in the exponent may be simplified as

$$\begin{aligned} & (y - \mu 1_{nr} - \mathbf{X}\beta) \, \Phi_1^{-1} \, (y - \mu 1_{nr} - \mathbf{X}\beta)' \\ = & \sum_{i=1}^{r}(y_i - \mu 1_{nr} - X\beta)' \, \frac{1}{\sigma_i^2} \, [I_n - \frac{\sigma_\alpha^2}{\sigma_i^2 + n\sigma_\alpha^2} \, J_n](y_i - \mu 1_{nr} - \mathbf{X}\beta) \\ = & \sum_i \frac{1}{\sigma_i^2} \sum_j (y_{ij} - \mu - x_j')^2 - \sum_i \frac{1}{\sigma_i^2} \frac{\sigma_\alpha^2}{\sigma_i^2 + n\sigma_\alpha^2} [\sum_j (y_{ij} - \mu - x_j' \, \beta)]^2 \, , \end{aligned}$$

where $x_j' = (x_{j1}, x_{j2} \ldots, x_{jk})$.

Thus the loglikelihood function may be written as

$$\begin{aligned} l_1 \;=\; & -\tfrac{nr}{2} \, ln(2\pi) - \tfrac{n-1}{2} \, \textstyle\sum_i \, ln \, \sigma_i^2 - \tfrac{1}{2} \textstyle\sum_i \, ln(\sigma_i^2 + n\sigma_\alpha^2) - \tfrac{1}{2} \textstyle\sum_i \textstyle\sum_j \, \tfrac{1}{\sigma_i^2} \, (y_{ij} - \mu - x_j'\beta)^2 \\ & + \tfrac{1}{2} \textstyle\sum_i \, \tfrac{1}{\sigma_i^2} \, \tfrac{\sigma_\alpha^2}{\sigma_i^2 + n\sigma_\alpha^2} \, [\textstyle\sum_j (y_{ij} - \mu - x_j' \, \beta)]^2 \end{aligned} \tag{6.7}$$

Partially differentiating l_1 with respect to the parameters and equating to zero, the ML estimating equations for the location parameters and the VCs may be obtained as

follows:

$$(n-1)(1-\frac{S_i^2}{\sigma_i^2})+\frac{(\sigma_\alpha^2}{\sigma_i^2+n\sigma_\alpha^2)^2}[(2+n\sigma_\alpha^2/\sigma_i^2)\ D_i^2+\frac{\sigma_i^2}{\sigma_\alpha^2}(\sigma_i^2+n\sigma_\alpha^2)] \ = \ 0\,,$$

$$\sum_i \frac{1}{\sigma_i^2+n\sigma_\alpha^2}\,[1-\frac{n}{\sigma_i^2+n\sigma_\alpha^2}\ D_i^2] \ = \ 0\,,$$

$$\sum_i \frac{1}{\sigma_i^2+n\sigma_\alpha^2}\sum_j (y_{ij}-\mu-x_j'\beta)\bar{x}_j' \ = \ 0\,,$$

$$\Rightarrow \sum_j (\bar{y}_{jw}-\mu-x_j'\beta)\bar{x}_j' \ = \ 0\,,$$

$$\sum_j (y_{jw}-\mu-x_j'\beta) \ = \ 0\,.$$

Here $S_i^2=(n-1)^{-1}\sum_j(y_{ij}-\mu-x_j'\beta)^2, D_i(\bar{y}_i-\mu-x_j'\beta), \bar{y}_i=\frac{1}{n}\sum_i y_{ij}$,

and $\bar{y}_{iw}=\sum_i w_i\,y_{ij}/\sum_i w_i$, where $W_i=\frac{1}{\sigma_i^2+n\sigma_\alpha^2}$ is the weight.

Note that from the last two equations, the estimates of β and μ may be obtained respectively. These estimators are same as the WLS estimators. here $\hat{\beta}=(X'X)^{-1}X'y_w$, where X is $(n\times k)$ matrix of regressors defined earlier and $y_w=(y_{1w}\ y_{2w}\ldots y_{nw})'$ is the weighted mean vector. The weighted mean for each j (sample) is calculated by taking weight $w_i=\frac{1}{\sigma_i^2+n\sigma_\alpha^2}$ *with* y_{ij}. We mention here that for unequal σ_i^2s, explicit expressions for σ_i^2 and σ_α^2 can not be obtained from the above equations. Also, the ML estimators obtained from the above system of equations for σ_i^2 and σ_α^2 may take negative values. If negative estimate of the variance component(s) is encountered, it may be replaced by zero and the iteration process may be continued.

6.5.2 Estimation of Random Component

In the model (6.3), the random component α_i represents the effect due to i^{th} Taster. This may be looked upon as the variation in the average score due to i^{th} assessor. This

unobservable random part can be estimated using the technique discussed in the fifth chapter.

The conditional distribution of α_i given the total errors $\alpha_i + u_{i1}, \alpha_i + u_{i2}, \ldots, \alpha_i + u_{in}$ can be obtained and an approximate summary measure of this conditional distribution may be assumed to represent the random component. The computational details follows from Section 5.5.3 of Chapter 5. The conditional probability of α_i given the total errors is

$$p(.) = p(\alpha_i \mid \alpha_i + u_{i1}, \alpha_i + u_{i2} \ldots \alpha_i + u_{in})$$

The conditional distribution being normal the mean of this conditional distribution may be assumed to give an approximate measure of α_i. From the above form we may write,

$$\hat{\alpha}_i = E(\alpha_i \mid \alpha_i + u_{i1}, \alpha_i + u_{i2} \ldots \alpha_i + u_{in}) = \frac{n\,\hat{\sigma_\alpha^2}}{\hat{\sigma_i^2} + n\,\hat{\sigma_\alpha^2}}(y_{i0} - x'_{ij}\hat{\beta})$$

6.5.3 Discussion on the Estimator of β

We note that $\hat{\beta}$ obtained here is nothing but the GLS or WLS estimator $\widehat{\beta_{GLS}} = (\mathbf{X'\Phi^{-1}X})^{-1}\mathbf{X'\Phi^{-1}y}$. This is obvious since under normality assumption ML and GLS estimates are same. It is also known that the covariance matrix of the estimated regression coefficients is $Var(\widehat{\beta_{GLS}}) = (\mathbf{X'\hat{\Phi}^{-1}X})^{-1}$. The estimated value of the covariance matrix is found by replacing the error variances by their ML estimators. Here $\hat{\beta}$ is consistent as well as efficient.

Again we may think of a transformation matrix P such that $P'P = \Phi^{-1}$. Then the estimator of β is $\hat{\beta} = (X'P'PX)^{-1}(X'P'Py) = [(PX)'PX]^{-1}[(PX)'Py] = (X^{*'}X^{*})^{-1}(X^{*'}y^{*})$, where X^{*} is the transformed design matrix and y^{*} is the transformed response vector. In this case we need not obtain the unknown Φ to estimate the fixed part of the model (6.5). Clearly $\widehat{\beta}_{GLS}$ reduces to $\widehat{\beta}_{OLS}$ under this transformation as the GLS estimator can be obtained applying OLS to the transformed observations (y^{*}, X^{*}).

If interest lies only with the estimation of β, then the problem boils down in choosing an appropriate form of the transformation matrix P only. Such a formulation would

substantially reduce the computation as we need not obtain Φ^{-1} in this case.

In many longitudinal analysis, the primary interest lies in making inference about β. When the mean response is of primary interest, we may use the WLS estimation technique. Here a symmetric weight matrix W may be introduced and $\hat{\beta}_w$ is obtained as $\hat{\beta}_w = (\mathbf{XWX})^{-1}\mathbf{X'Wy}$. This estimate is unbiased whatever may be the choice of W. If $W = I$, we obtain OLS estimator. We note here that under random effects formulation, the GLS estimator of β based on the true VCs are BLUE. All the feasible GLS estimators considered are asymptotically efficient as neither n or r tends ∞. Maddlla and Mount (1973), Taylor (1980) and Baltagi (1981a) conducted Monte Carlo studies and found little to choose among various feasible GLS estimators in small samples and argued in favour of ANOVA feasible GLS and ML to ensure that these do not yield drastically different results.

6.5.4 The REML Estimation for Model I

In a homoscedastic situation the ML estimator of VC is not unbiased. One needs to adjust for its degrees of freedom. This is overcome by REML approach. In this section we obtain the REML estimators of variance components for the linear error component model - Model I.

In case of general linear model $y \sim N(\mathbf{X}\beta, \mathbf{\Phi})$, the REML estimator may be defined as a ML estimator based on a linearly transformed set of data $y^* = Ay$ such that the distribution of y^* does not depend on β. One way of achieving this is by taking A to be the matrix which converts y to OLS residuals.

$$A = I - X(X'X)^{-1}X$$

Here we have $y_j^* = y_j - x_j' \hat{\beta}_{OLS}$, where $\hat{\beta}_{OLS} = (X'X)^{-1} X'y$. It follows that $E(y_j^*) = 0$ or $E(y^*) = 0$ for any choice of β, and in fact, the distribution of y^* is independent of β. Under this transformation the reduced profile loglikelihood can be shown to be

$$\hat{l}_R = -\frac{1}{2}[(nr - k)\ \ln(2\pi) + \ln\ |\ \Phi\ | + \ln\ |\ X'\ \Phi^{-1}\ X\ | + (y - X\hat{\beta})'\ \Phi^{-1}(y - X\hat{\beta})]$$

where the MLE of the parameters are obtained by maximizing

$$l_{ml} = -\frac{1}{2}[nr\ \ln(2\pi) + \ln\ \mid \Phi + \mid (y - X\hat{\beta})'\ \Phi^{-1}(y - X\hat{\beta})]$$

It follows from the results that the general linear model incorporates only a simple modification of the ML algorithm in the earlier section.

The restricted likelihood function for the Model I can be maximized by using the EM algorithm or the Newton-Raphson algorithm. Various packages now exist to maximize or minimize the loglikelihood function (for example, SPLUS, using *nlmin* command).

If we compare the l_{ml} and l_R, we observe that the basic difference lies with the term $\frac{1}{2}\ \ln\ \mid X'\Phi^{-1}X\ \mid$ apart from the coefficient of $\ln(2\pi)$. It may be noted that $X'\Phi^{-1}X$ is a $(k \times nr) \times (nr \times nr) \times (nr \times k) = (k \times k)$ matrix. So this term is typically of order k, where as l_{ml} is of order nr suggesting, correctly, that the distinction between ML and REML estimation is important only when k is relatively large. Many authors have discussed the relative merits of ML and REML estimators for covariance parameters. Cullins et al. (1990), Verbyla et al. (1990) apply REML in longitudinal data settings, whilst Tunnichiffe-Wilson (1989) uses it for time-series estimation. One of Tunnichiffe-Wilson examples shows how REML copes much more efficiently with a near singular variance matrix than does ML estimation. The two methods are asymptotically equivalent as either or both n and r tend to infinity for fixed k. When k tends to infinity, comparisons unequivocally favour REML. In summary, ML and REML estimators will often give very similar results. But when they differ substantially, the REML estimators are preferable. But we note here that not much is known about this while dealing with small sample size (fixed samples).

6.6 ML Estimation for Model II

In the heteroscedastic Model II we assume that $\alpha_i \sim iid(0, \sigma_i^2)$ and $u_{ij} \sim iid(0, \sigma_u^2)$. The random component α_i is assumed to be independent of u_{ij} and $\alpha_l (i \neq l)$. In this case, for the model (6.5) we have $E(y_{ij}/x_{ij}) = \mu + x'_{ij}\beta$, $Var(y_{ij}/x_{ij}) = \sigma_i^2 + \sigma_u^2$, and $Cov(y_{ij}, y_{ik}) = \sigma_u^2$ for $j \neq k$.

The variance-covariance matrix for this model may be obtained as (see Section 5.6.2 of Chapter 5)

$$\begin{aligned} E(e\,e') &= \Phi_2 = Z\,\Sigma\,Z' + \sigma_u^2\,I_{nr} \\ &= \Sigma \otimes J_n + \sigma_u^2(I_r \otimes I_n) \\ &= diag(\sigma_1^2 J_n + \sigma_u^2 I_n \;\; \sigma_2^2 J_n + \sigma_u^2 I_n \ldots \sigma_r^2 J_n + \sigma_u^2 I_n) \\ &= diag(S_1, S_2, \ldots S_r)\,, \end{aligned}$$

where $\Sigma = diag(\sigma_1^2, \sigma_2^2, \ldots \sigma_r^2)$.

The determinant and inverse of Φ_2 may be obtained as $\mid \Phi \mid = \mid S_1 \mid\mid S_2 \mid \ldots \mid S_r \mid = (\sigma_u^2)^{r(n-1)} \prod_{i=1}^{r} (\sigma_u^2 + n\sigma_i^2)$, and $S^{-1} = \frac{1}{\sigma_u^2} I_n - \frac{\sigma_i^2}{\sigma_u^2(\sigma_u^2 + n\sigma_i^2)} J_n$.

Now the likelihood function for the Model II may be written as (under the gaussian assumption)

$$L_2 = (\sqrt{2\pi})^{-nr}(\sigma)^{-r(n-1)}[\Pi_i(\sigma_u^2 + n\sigma_i^2)]^{-\frac{1}{2}}\, exp[-\frac{1}{2}(y - \mu - \mathbf{X}\beta)'\,\Phi_2^{-1}\,(y - \mu - \mathbf{X}\beta)\,,$$

and the quadratic part in the exponent may be simplified as

$$\sum_{i=1}^{r}(y_i - \mu - X_i\beta)' \left[\frac{1}{\sigma_u^2} I_n - \frac{\sigma_i^2}{\sigma_u^2(\sigma_n^2 + n\sigma_i^2)} J_n\right] (y_i - \mu - X_i\beta)$$

$$\begin{aligned} &= \frac{1}{\sigma_u^2}\sum_i (y_i - \mu - X_i\beta)'(y_i - \mu - X_i\beta) \\ &\quad - \sum_i \frac{\sigma_i^2}{\sigma_u^2(\sigma_n^2 + n\sigma_i^2)}(y_i - \mu - X_i\beta)'\, J_n\, (y_i - \mu - X_i\beta) \end{aligned}$$

$$= \frac{1}{\sigma_u^2}\sum_i\sum_j (y_{ij} - \mu - x_j'\beta)^2 - \sum_i \frac{\sigma_i^2}{\sigma_u^2(\sigma_u^2 + n\sigma_i^2)}\left[\sum_j (y_{ij} - \mu - x_j'\beta)\right]^2$$

where $x_j' = (x_{j1}, x_{j2}, \ldots x_{jk})$.

The loglikelihood function for the Model II may be written as

$$l_2 = -\frac{nr}{2}\ln(2\pi) - \frac{r(n-1)}{2}\ln\sigma_u^2 - \frac{1}{2}\sum_i \ln(\sigma_u^2 + n\sigma_i^2)$$

$$-\frac{n-1}{2\sigma_u^2}\sum_i S_i^2+\frac{1}{2}\sum_i \frac{\sigma_i^2}{\sigma_u^2(\sigma_u^2+n\sigma_i^2)}D_i^2\,, \qquad (6.8)$$

where $(n-1)\,S_i^2=\sum_{j=1}^{n}(y_{ij}-\mu-X_j'\beta)^2$ and $D_i=\sum_j(y_{ij}-\mu-x_j'\beta)$.

Partially differentiating (6.8) with respect to $\sigma_u^2,\sigma_i^2,\mu$ and β, we obtain the following ML estimating equations :

$$\sigma_u^2+n\sigma_i^2=\tfrac{1}{n}\left[\textstyle\sum_j(y_{ij}-\mu-x_j'\beta)\right]^2$$

$$\Rightarrow\quad \sigma_i^2=\left[\tfrac{1}{n}\textstyle\sum_j(y_{ij}-\mu-x_j'\beta)\right]^2-\tfrac{1}{n}\sigma_u^2 \qquad (6.9)$$

$$\Rightarrow\quad \sigma_i^2=D_i^2-\tfrac{1}{n}\sigma_u^2$$

$$(n-1)\left(r-\sum_i\frac{S_i^2}{\sigma_u^2}\right)+r+\sum_i\frac{n\sigma_i^2}{\sigma_u^2}=0$$

$$\sum_i\frac{1}{\sigma_u^2+n\sigma_i^2}\sum_j(y_{ij}-\mu-x_j'\beta)\,\bar{x}_j^1=0$$

$$\Rightarrow\quad \sum_j(y_{iw}-\mu-x_j'\beta)\,\bar{x}_j^1=0\quad\text{and}$$

$$\sum_j(\bar{y}_{iw}-\mu-x_j'\beta)=0$$

$$\Rightarrow\quad \mu=\frac{1}{n}\sum_j(\bar{y}_{jw}-x_j'\beta)\,,$$

where $\bar{y}_{jw}=\sum_i W_i y_{ij}/\sum_i W_i$, $W_i=\frac{1}{\sigma_u^2+n\sigma_i^2}$. Here $\bar{y}_{jw}$ is the weighted mean for j^{th} sample, where weights are obtained from the estimates of variance components. Clearly the ML estimators of β and μ are the weighted least square estimators.

6.7 Residual Analysis for the Error Component Model With General Covariance Structure

Haslett and Hayes (1998) has discussed the general theory for residuals from the general linear model. For the model $y = \mu + X\beta + Z\alpha + u$, we have $y \sim (X\beta, \Phi)$, where $\Phi = E(WW')$ for $W = Z\alpha + u$. The GLS estimator of β is $\widehat{\beta} = (X'\ \Phi^{-1}\ X)^{-1}\ X'\ \Phi^{-1}y$ with $Var(\widehat{\beta}) = (X'\ \Phi^{-1}\ X)^{-1}$. Here the classical residual may be defined as $\widehat{u} = y - X\widehat{\beta} = \Phi QY$, where $Q = \Phi^{-1}\ P$ and $P = \Phi^{-1}X(X'\Phi^{-1}X)^{-1}\ X'\ \Phi^{-1}$. Also, $Var(\widehat{u}) = \Phi Q\Phi = G$.

The 'lack of fit' statistic for the model is defined as $S = \widehat{u}'\ \Phi^{-1}\ \widehat{u} = y'Qy$. Note here that to estimates S, we may use the ML estimators of variance components. The closer the values of S to zero, the better is the fit.

We note here that since the method assumes normality, there is also a need to assess the adequacy of that assumption. For RE models, Ryan and Dempster (1984) proposed the weighted normal plot as a graphical approach to the assessment of normality. However, failure of approximate normality assumption does not invalidate the estimates of location parameters since weighted least squares estimates are unbiased and consistent under very broad conditions. It does, however, invalidate the usual tests and confidence intervals based on normality theory.

The 'bootstrap' provides an alternative approach to assessing the distributional properties of estimators without reliance on normality assumptions. The 'bootstrap' approach provides an estimate of the sampling distribution of an estimator based only on that estimation procedure and the sample in hand, without appealing to distributional assumptions. If an approximately optimal estimator can be defined, the bootstrap can be used to assess its properties.

6.8 Study on Improvement of Quality of CTC Teas Through Process Modification

A new process of manufacture of CTC tea has been developed at Tocklai Experimental Station, by modifying the sequences of manufacturing steps. In the modified method, the plucked shoots were withered, rolled, rotorvened, fermented and then taken to a CTC machine and dried instead of fermenting after CTC cut. The final products obtained from the new sequences furnished marked improvement in shelf-life of the commercially manufactured teas. This modified manufacturing process has, therefore, been adapted by different commercial gardens of Assam and Dooars regions of India and outstanding results have been observed in terms of auction price realization.

Shoots plucked from Tocklai experimental plot were manufactured at the Miniature factory. The whole lot of leafs were divided into two equals parts. One part received the conventional manufacturing process, whereas the other part received the innovative modified manufacturing system. Thus we have the 'central' and 'experimental' samples. The biochemical parameters measured are moisture (MO), TF, TR, C, brightness (B), total colour (TC) and water-soluble solids (WSS). Detail technical discussion on the manufacturing methods and the chemical analysis is available in Pal et. al. (2000).

Four experienced Tasters have evaluated the samples in terms of 'valuation' on a 0-10 point structured scale. For each sample (control and experimental), each Taster made 10 repeats. But the replicated scores are not available, the mean of replicated scores on each sample is provided.

The whole batch of samples are divided into 'central' and 'experimental' group. In such a situation, the natural interest is to study whether the quality for given set of samples has really improved over the experimental condition or not. This is a 'too treatment group' situation and model has to be developed taking account of the two groups. We adjust our model 6.6 introducing dummy variable with the intercept term in the model matrix. We assume that only the intercept term varies between the groups, as separate slopes (β) is not suggestive for the given data set under study. We have examined the difference of β

coefficients two groups for each biochemical parameter. Thus we can write our model as

$$Y = \begin{pmatrix} Y^{(1)} \\ Y^{(2)} \end{pmatrix} = \begin{pmatrix} 1_m & 0 & X^{(1)} \\ 1_m & 1_m & X^{(2)} \end{pmatrix} \begin{pmatrix} \mu^* \\ b^* \\ \beta \end{pmatrix} + \begin{pmatrix} Z^{(1)} & 0 \\ 0 & Z^{(2)} \end{pmatrix} \begin{pmatrix} \alpha^{(1)} \\ \alpha^{(2)} \end{pmatrix} + \begin{pmatrix} e^{(2)} \\ e^{(2)} \end{pmatrix}$$

where $y^{(1)}$ and $Y^{(2)}$ and $(m \times 1)$ vectors of repeated response for experimental and control samples respectively, with $2m = n$. μ^* is the intercept and b^* is the dummy coefficient associated with the dummy variable. Similarly we have the partition of incidence matrices $Z^{(1)}$ and $Z^{(2)}$ and the random component. Here $\beta = (\beta_1\ \beta_2\ \ldots\ \beta_k)'$ is the vector of fixed coefficients associated with the regressors.

Following the loglikelihood function is (6.7) we may write the loglikelihood for the 'two-group' model as

$$\begin{aligned} \ln L \;=\; & -\frac{nr}{2}\ln(2\pi) - \frac{n-1}{2}\sum_i \ln\sigma_i^2 - \frac{1}{2}\sum_i ln(\sigma_i^2 + \sigma_\alpha^2) \\ & -\frac{1}{2}\sum_i \sum_{j=1}^{m} \frac{1}{\sigma_i^2}(y_{ij} - \mu^* - x_j^{(1')}\,\beta)^2 - \frac{1}{2}\sum_i \sum_{j=m+1}^{n} \frac{1}{\sigma_i^2}(y_{ij} - \mu^* - x_j^{(2')}\,\beta)^2 \\ & +\frac{1}{2}\sum_i \frac{1}{\sigma_i^2}\,\frac{\sigma_\alpha^2}{\sigma_i^2 + n\,\sigma_\alpha^2}\left\{\left[\sum_{j=1}^{m}(y_{ij} - \mu^* - b^* - x_j^{(1')}\beta)\right]^2\right. \\ & \left. +\left[\sum_{j=m+1}^{n}(y_{ij} - \mu^* - b^* - x_j^{(1')}\beta)\right]^2\right\} \end{aligned}$$

The ML estimators of parameters can be obtained by maximizing the loglikelihood function. We have obtained the estimators of the variance components and the regression coefficients along with the estimate of dummy coefficient using the **nlmin** command in the SPLUS computational package.

We note here that the estimators of variance components, dummy coefficient and the regression coefficients may be obtained similarly using the Model II and its corresponding likelihood function (6.12).

6.8.1 Data Analysis and Discussion

The measurements of biochemical parameters and the Tasters' scores for CTC tea samples are presented in Table 6.9 and Table 6.8 respectively. The basic statistics for all the chemical parameters and sensory scores, for control and experimental samples separately, have already been presented in the introductory chapter. The significance of difference between the mean levels over the two groups of samples for each chemical parameter have been tested. The mean levels of TR, TC and WSS differ significantly at 5% level (5% probability values of t with 16 d.f. is 2.12). Also each Taster's mean scores differ significantly over the control and experimental samples. It may be observed that the average scores of all the assessors are higher for the experimental samples as compared to those for control samples.

It may be observed from the basic statistics that the ranges of TF for control and experimental samples are (1.14m 1.60) and (1.17, 2.02) respectively; and those for TR are (13.04, 14.08) and (10.90, 13.72). The ranges of these two variables are quite large for the experimental samples. Similar is the case with WSS as well. The within sample variability is higher for TR in case of experimental samples (S.D. = 0.76) though that for TF is relatively low (S.D. = 0.24). Also, it is interesting to observe that the correlation coefficient between TF and TR is negative for experimental samples (-0.26) but that for the control samples is positive (0.15) although the degree of association is very low.

We now associate the biochemical quality parameters with the average (arithmetic mean) of four Tasters' scores. Dummy variable is introduced (1 for experimental and 0 for control) with the regressors. This is done to study whether the quality has really improved over experimental condition or not. The regression result is presented in Table 6.6. The fit is very poor with adjusted $R^2 = 0.26$ and SSE = 7.9549. All the regression coefficients along with the dummy coefficients are statistically insignificant.

The estimates of regression parameters and variance components using our error component regression model with dummy variable are presented in Table 6.7. To obtain the iterative estimates of implicit parameters we set the convergence tolerance at 10^{-5}. The

estimate of dummy coefficient b^* is positive and statistically highly significant (t-ratio : 6.84). Thus we may infer that the cup quality has significantly improved for the experimental samples, so far the given CTC samples are concerned. In fact, the experimental samples have received higher scoring in most of the cases. The estimated coefficients for both TF and TR along with WSS are statistically significant at 5% level. Interestingly, the coefficient estimates for MO and TC are negative, though statistically insignificant. The estimated scores for 'valuation' are given in the last column of Table 6.8. The estimated true scores are higher for the experimental samples in comparison to the control samples.

The variance component σ^2 represents the variation in average levels among the Tasters. A negligibly small value of σ^2 (= 0.0001) indicates the agreement among all the four Tasters in their average valuation for the given CTC samples. Thus we may say that the average scores given by Tasters for the CTC tea samples are nearly equal. On the other hand, σ_i^2s represents the within Taster variation. We note how that the unobservable regression error variance is confounded in σ_i^2 and can not be separated. As may be observed from Table 6.7, the estimated error variance is minimum for the Taster 2 and is maximum for Taster 1. These error variances may be considered as grinding factors to assess the reliability of a particular Taster's choice. Here σ_i^2 concerns precision. We could apply the LR test to study the significance of difference between σ_i^2s. But we restrict to do so keeping in view the small sample size.

For residual analysis, the scatter plot of error against the estimated response was studied and an approximate normality was ascertained from the plot. The estimated value of SSE is 0.3639, which is much less than the SSE obtained from the regression on average. Also the value of estimated loglikelihood is much higher in case of error component model as compared to the regression on average. In fact, the difference between the values of $ln\ L$ is 27.76. We can apply the information criteria based test (Vonesh and Chinchilli, 1997) discussed in the second chapter, to compare the fit. We note here that the general linear regression model, where the response represents average of four Tasters' scores, and the error-component regression model are non-nested models. The Akaike; information

criterion (AIC) values for the regression on mean and the error-component model are -27.1917 and -22.5782. Since the AIC value is larger in case of error component model, we conclude that the fit is better with the 'two-group' variant of Model I.

Also note that as the estimated value of σ^2 is very close to zero, supporting the model assumption of constant variability among the average levels of Tasters' scores, we do not try the 'two-group' variant of model II.

We note here that some statistical confusions may arise while dealing with small sample size, as it is our case. One may raise arguments on the danger of incorrect conclusion about the slope coefficients, while dealing with small samples. But we note here that, though it appears as if we have 18 samples (two groups combined), we have practically 72 observations because of 4 repeats on each sample for the response variable. Under the assumption of same β coefficients for both groups, we are estimating only 13 parameters, namely 5 variance components, 6 slope coefficients, intercept and dummy. Thus we have 59 degrees of freedom left for estimating the β coefficients. In this situation we do not think that there is any real danger of incorrect conclusion about the slope coefficients. The real danger due to small samples of this type may come from the problem of multicollinearity since the biochemical parameters (regressors) are not repeatedly observed.

We restrict ourselves to draw general conclusion about the statistical significance of the biochemical parameters, as we believe, a more planned experiment is required with strong data base to decide upon the significance of biochemical parameters in this modified CTC manufacturing process. Experiments should be carried out at different CTC tea producing regions and other relevant chemical parameters should be measured. Assessment of individual theaflavins, which have different astringencies, can be made by HPLC, but is difficult to undertake such assessment at production level. For the given CTC samples under study, only TF, TR and WSS are statistically significant in explaining the overall valuation. Note that the significance is specific to the quality attribute 'V', and should not be generalized for other attributes like strength, briskness, etc.

At commercial gardens, the manufactured teas are classified in terms of different grades like Brokens, Fannings, Dust, etc. The higher the percentage of Brokens, the higher would

be the price for a particular tea lot. For obvious reasons, the producers would favour a manufacturing process which gives higher Broken percentage for a given lot of tea. The modified manufacturing process was tried by three gardens of Dooars region and four gardens of Assam region of India. The grade percentage obtained for both control and experimental teas are presented in Table 6.5. As many be seen from the Table, Broken percentages for experimental samples are much higher in all the commercial gardens.

The Tasters and the manufacturers have observed that the tea samples obtained by the modified manufacturing process has better shelf-life and fetched higher prices. Indeed, the teas are brighter and brisker. The price realization for these teas have been found to be the highest amongst the best CTC category in both Calcutta and Guwahati Auction Centers.

6.8.2 Analysis of Data Set 7

As discussed in the first chapter, for a set of 18 CTC samples, a panel of 4 Tasters have given scores (on an uniform structured scale) on 'strength' and 'quality' attributes separately and independently. The biochemical parameters measured are TF, TR, WSS, TC and C.

We first test the significance of difference among the mean scores of four Tasters for strength and quality. That is, we text if the profiles for means of Tasters' scores are parallel and coincide. This text is discussed in the Chapter 2. The estimated value of T^2 is 6.43 and the 5% critical value of F with 1 and 72 degrees of freedom is . Thus we may reject the hypothesis of parallel and coincident mean profiles at 5% level of significance. This implies that the Tasters differ in their average choices on strength and quality for the given set of 18 CTC tea samples.

In the second stage of our initial diagnostics, we calculated the F-statistics (discussed in Section 6.2) to study the stability of location parameters, before we pool the data for four Tasters' scores. We calculate the F-values separately for strength and quality. The calculated values of F-statistic for strength and quality are 2·13 and 1·97 respectively. The hypothesis of stable estimates of regression coefficients is accepted, as the critical

value of F with d.f. at 5% level is On the basis of this test result, we may use our error component regression model with same β coefficients for all the individual responses.

We now adopt the technique of regression on means of Tasters' scores, discussed in the introductory section. The results of multiple regression analysis (where response variable represents mean of 4 Tasters' scores) are presented in Table 6.6 and Table 6.6, for strength and quality respectively. The fits are very poor in both the cases with very low values of adjusted R^2 and high values of SSE. All the chemical parameters are statistically insignificant in explaining the quality, as evident from the t-ratios. Only the Caffeine (C) has come out significant at 5% level in explaining the strength.

The association of chemical parameters with the scores on strength indicates that only TF and C have significant influence on the attribute strength. Both ML and REML estimates provide the same indication. The coefficient estimates of TR and WSS are negative, though statistically insignificant. Again for the attribute 'quality', the parameters TF, TR, TC and C have come out significant. It is interesting to observe that $\widehat{\beta}_{TR}$ is positive now. We have observed the similar behaviour of TR in Chapter 3 also. The significant influence of TF and TR on 'quality' in this case support the common belief among the biochemists about the importance of TF and TR in the CTC tea quality assessment (Yumanishi, 1995, and the references therein).

The error variance (σ_i^2) is highest for Taster 1 and is lowest for the Taster 2, in case of strength. But in case of quality, though the variation is highest for the first Taster, it is lowest for the Taster 4. Thus the pattern of error variances is not same for the two different quality attributes. The variation among the average levels of Tasters (σ^2) is not at all close to zero. It means that the Tasters do not agree on average, for both strength and quality, for the given set of CTC samples. Also the within sample variability is high for all the Tasters, as evident from the estimates of σ_i^2s.

The fit is far better with the heteroscedastic error component model as compared to the regression on average. The $ln\ L$ values are much higher and the SSE values are very small. We note here that the performance of ML and REML estimation is almost equivalent.

Comparing the $ln\ L$ and SSE values we can not claim any significant improvement with REML estimation. We have studied the patters of residuals in both the cases. The scatter plots of residuals against the estimated response exhibit similar patterns. However, we do not present the scatter plots here.

We note here that we have also tried the Model II (both ML and REML) on the strength and quality data separately. The estimated values of loglikelihood (ML) for strength and quality data re -5.1340 and -5.0698; and given the equal number of estimated parameters in both the cases, the AIC values are certainly larger for Model I. It means that though the estimated values of σ^2s are not very close to zero (Table 6.8), still the use of Model I remain valid and gives a better first in comparison to Model II. That is why we prefer to present the estimates obtained using Model I for given Data Set.

As the regression on mean of Tasters' scores do not provide a good fit in our pursuit to associate the chemical information with sensory evaluation, we proceed to fit the error component model. Here we prefer to use the Model II as in initial diagnostics we have observed that the average scores of the four Tasters for the given sample differ significantly. The ML and REML estimates of β coefficients along with the estimator of VCs, separately for strength and quality, are presented in Table 6.8.

Table 6.1 : OLS Regression on average of four Tasters

Var	Estimate	S.E.	t-ratio
$\widehat{\mu}^*$	-7.9738	23.13	0.3447
$\widehat{b}^*$	0.9166	0.6872	1.344
$\widehat{\beta}_{MO}$	-0.5616	0.6992	0.738
$\widehat{\beta}_{TF}$	1.3519	1.245	1.086
$\widehat{\beta}_{TR}$	-0.3516	0.4071	0.8638
$\widehat{\beta}_{TC}$	-1.046	1.26	0.83
$\widehat{\beta}_{WSS}$	0.4159	0.5263	0.7903
$\widehat{\beta}_C$	0.2229	0.5846	0.3813
$\bar{R}^2$	0.26		
SEE	7.9549		
$ln\ L$	-18.1917		

Table 6.2 : ML estimates of parameters and variance components

Parameter	Estimates	SE
$\widehat{\sigma}_1^2$	1.0738	
$\widehat{\sigma}_2^2$	0.3639	
$\widehat{\sigma}_3^2$	1.0273	
$\widehat{\sigma}_4^2$	0.9090	
$\widehat{\sigma}^2$	0.0001	
$\widehat{\mu}^*$	8.2874*	2.7738
$\widehat{b}^*$	10.2844*	1.5028
$\widehat{\beta}_{MO}$	-0.1451	0.1988
$\widehat{\beta}_{TF}$	1.2477*	0.5842
$\widehat{\beta}_{TR}$	-0.0046*	0.0021
$\widehat{\beta}_{TC}$	-0.8047*	0.5907
$\widehat{\beta}_{WSS}$	0.0731*	0.0245
$\widehat{\beta}_C$	0.4353	0.2735
$ln\ L$	9.5682	
SEE	0.3639	

*indicates significance at 5% level.

Table 6.3 : Tasters' scores on 'valuation' and the 'estimated scores'

Sample	Taster (1)	Taster (2)	Taster (3)	Taster (4)	Estimated Score
Exp.	6	8	7	9	8.1684
Cont.	6	7	6	7	6.7345
Exp.	7	9	9	8	8.2174
Cont.	7	6	7	6	6.6817
Exp.	7	9	8	8	7.8868
Cont.	7	7	6	6	6.7421
Exp.	8	9	9	9	8.8253
Cont.	8	8	8	5	7.2761
Exp.	8	8	8	9	8.0479
Cont.	7	6	8	7	6.7921
Exp.	8	9	8	10	8.9594
Cont.	7	7	8	8	6.7714
Exp.	7	8	7	9	7.6603
Cont.	4	8	6	7	6.6665
Exp.	9	8	7	8	8.0725
Cont.	6	6	8	6	6.5302
Exp.	9	8	8	10	8.4187
Cont.	6	7	4	8	6.6053

Table 6.4 : Measurements on biochemical parameters

Sample	MO	TF	TR	TC	WSS	C
Exp.	3.48	1.75	11.26	5.69	45.73	4.07
Cont.	3.72	1.40	13.14	5.17	44.12	3.92
Exp.	3.05	1.80	11.78	5.79	45.86	4.70
Cont.	3.16	1.46	13.12	5.38	44.40	4.30
Exp.	3.36	1.70	11.63	5.75	45.72	4.68
Cont.	3.50	1.45	13.24	5.30	44.53	4.25
Exp.	2.97	1.17	11.27	5.70	44.65	4.06
Cont.	2.85	1.14	13.04	5.24	44.00	4.00
Exp.	3.12	1.50	11.72	5.98	44.91	4.18
Cont.	2.45	1.46	13.09	5.10	43.21	4.14
Exp.	3.86	1.79	12.09	5.85	46.57	5.67
Cont.	4.03	1.45	13.12	5.45	45.49	4.66
Exp.	3.59	2.02	10.90	5.45	44.90	3.73
Cont.	3.29	1.60	14.08	5.05	43.95	3.67
Exp.	3.98	1.46	13.72	5.89	44.63	3.97
Cont.	4.10	1.56	13.18	5.50	45.52	4.65
Exp.	3.65	1.79	12.06	5.77	46.58	5.37
Cont.	4.00	1.53	13.15	5.51	45.55	4.66

Table 6.5 : Grade percentages in different gardens of Assam and Dooars region

Gardens	Brokens		Fannings		Dust		Secondary	
	Cont.	Exp.	Cont.	Exp.	Cont.	Exp.	Cont.	Exp.
Garden A	63.9	67	16.4	16.2	4.4	3.4	15.3	13.4
Garden B	68.3	74.7	12.4	11.7	6.7	6.0	12.6	7.6
Garden C	70.6	78.1	16.9	12.3	12.5	9.6	-	-
Garden D	36.4	47.19	23.37	23.62	29.65	29.19	10.58	-
Garden E	38.45	60.40	18.01	22.29	26.05	13.38	17.49	3.93
Garden F	42.28	57.05	24.05	21.35	19.17	13.60	14.40	8.00
Garden G	60.01	69.50	14.58	13.33	15.40	9.16	10.01	8.01

* Gardens A, B and C belong to Dooars region and the rest gardens belong to the Assam region.

Table 6.6 : Regression results on average of Tasters' score on strength

Parameter	Estimate	S.E.	t-ratio
β_0	-7.9046	4.4185	1.7889
β_{TF}	0.37846	0.3375	1.12
β_{TR}	-0.2457	0.2102	1.1688
β_{WSS}	-0.1794	0.3158	0.568
β_{TC}	0.789	0.562	1.4039
β_C	-0.4873	0.2296	2.1228
$\bar{R}^2$	0.38		
SSE	4.7346		
$ln\ L$	-14.8560		
$\chi^2_{(normal)}$			

Table 6.7 : Regression results on average of Tasters' score on Quality

Parameter	Estimate	S.E.	t-ratio
β_0	6.7463	11.27	0.5986
β_{TF}	0.4333	1.125	0.3853
β_{TR}	0.2042	0.2681	0.7617
β_{WSS}	-0.2274	0.3451	0.6591
β_{TC}	1.5301	1.042	1.469
β_C	+0.3418	0.7211	0.4739
$\bar{R}^2$	0.1938		
SSE	7.6734		
$ln\ L$	-17.8673		
$\chi^2_{(normal)}$			

Table 6.8 : ML and REML estimates of β and VC for strength and quality

Parameter	Strength		Quality	
	ML	REML	ML	REML
$\hat{\beta}_0$	12.2921	9.4521	24.3666	25.56
	(11.9579)	(10.234)	(10.7482)*	(10.35)*
$\hat{\beta}_{TF}$	0.2697	0.5621	0.1821	0.5642
	(0.1172)*	(0.1103)*	(0.0074)*	(0.131)*
$\hat{\beta}_{TR}$	-0.1985	-0.2356	1.5865	2.0321
	(0.6782)	(0.1232)*	(0.6166)*	(0.132)*
$\hat{\beta}_{WSS}$	-0.0906	-0.2019	0.0834	0.2879
	(0.2175)	(0.2356)	(0.1643)	(0.3121)
$\hat{\beta}_{TC}$	0.1913	0.3256	0.3870	0.9875
	(0.2225)	(0.2545)	(0.1879)*	(0.1268)*
$\hat{\beta}_C$	0.3093	0.4352	0.3973	0.6546
	(0.2326)*	(0.1527)*	(0.20)*	(0.12)*
$\hat{\sigma}_1^2$	3.8936	3.97	4.1239	3.9478
$\hat{\sigma}_2^2$	1.6936	1.5987	1.5523	2.0218
$\hat{\sigma}_3^2$	2.8236	2.3789	3.0902	3.4635
$\hat{\sigma}_4^2$	2.8249	2.4563	1.0943	1.4849
$\hat{\sigma}^2$	0.98	0.9890	0.974	0.9231
$ln\ L$	-3.97	-3.002	-3.51	-3.01
SSE	0.9984	0.8445	0.7644	0.7342

Table 6.9 : Estimated true scores on strength and quality

Strength		Quality	
ML	REML	ML	REML
6.2873	6.2312	6.4998	6.5535
6.3595	6.3989	6.6138	6.6336
5.0065	5.0787	4.7579	4.9968
6.6794	6.9872	6.8622	6.8877
6.2886	6.3254	6.9917	7.0120
5.9389	5.9987	6.2629	6.3636
6.0122	6.1210	6.1783	6.4451
5.9405	5.9868	6.1093	6.1363
6.4861	6.5423	6.1390	6.3839
5.8548	5.9851	6.1204	6.3135
6.0705	6.1310	5.9473	5.9983
6.0712	6.0909	5.6493	5.8871
6.6923	6.8878	6.6592	6.6979
6.3714	6.3939	6.8803	6.9791
6.0323	6.1312	5.8591	5.8787
6.6723	6.6961	6.8507	7.0012
6.2106	6.3223	5.5057	5.6561
6.3688	6.3698	6.0501	6.1535

Table 6.10 : Measurements on biochemical parameters

TF	TR	WSS	TC	C
1.40	13.14	44.12	5.17	3.92
1.75	11.26	45.73	5.69	4.07
1.46	13.12	44.40	5.38	4.30
1.80	11.78	45.86	5.79	4.70
1.45	13.24	44.53	5.30	4.25
1.50	13.59	45.72	5.73	4.17
1.14	13.63	44.00	5.24	4.00
1.17	11.27	44.65	5.70	4.06
1.46	13.09	43.20	5.10	4.14
1.50	11.72	44.91	5.98	4.18
1.45	13.12	45.49	5.45	4.66
1.37	14.00	46.57	5.85	4.52
1.60	14.08	43.95	5.05	3.67
2.02	10.09	44.90	5.45	3.73
1.56	13.18	45.52	5.50	4.65
1.46	13.72	44.63	5.89	3.97
1.53	13.15	44.55	5.51	4.66
1.79	12.06	46.58	5.77	4.37

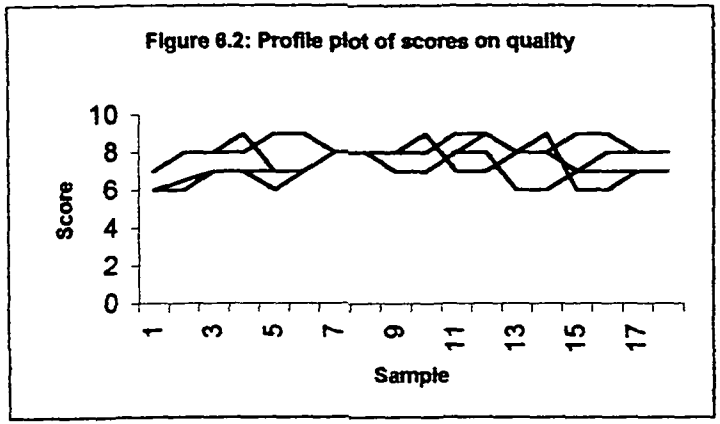


Figure 6.2: Profile plot of scores on quality
Score
10
8
6
4
2
0
1
3
5
7
9
11
13
15
17
Sample

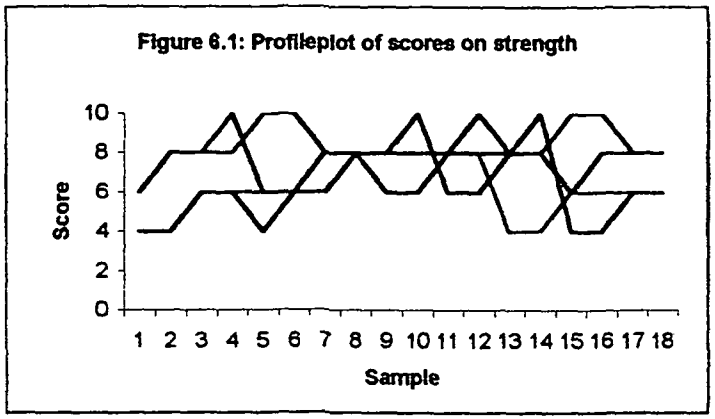


Figure 6.1: Profileplot of scores on strength
Score
10
8
6
4
2
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
Sample

CHAPTER 7

TWO-WAY ERROR COMPONENT REGRESSION MODELS WITH HETEROSCEDASTIC ERROR

7.1 INTRODUCTION

As an extension of the one-way error component model, we study the two-way error component regression models in this chapter. The error component is decomposed into two parts to incorporate the sample specific effects apart from the effects due to Tasters. This is done to check for the variation over samples.

Consider the Data Set 8, where 30 samples represent different CTC clones. Single sample for each clone has been observed. The chemical characteristics may differ over clones. This difference is also likely to influence the overall quality. Thus apart from studying the Tasters' variation, it is also important to consider the clonal variation as well. We may legitimately assume that the chemical characteristics is fixed specific to a particular CTC clone. Thus in regression setup, we may assume the sample specific effects to be fixed. The variation due to Tasters may be assumed random. Thus we have a mixed effects error components regression model.

The two-way error component regression models are well developed in the statistical literature and a detail account of such formulations is due to Baltagi (1995). The fixed and random effects models with homoscedastic error components have been discussed. The GLS and ML estimators of the location parameters are developed along with the testing procedures about the fixed effects. We simply concentrate on the mixed effects model with heteroscedastic formulation of variance-covariance matrix. The ML estimators of the location parameters along with the estimators of variance components are discussed. The possible heteroscedastic formulations and the estimation procedures are discussed in section 7.3. The basic model is presented in section 7.2. The section 7.4 contains the

results obtained using Data Set 8 and discussions. The original row data on Tasters' scores and five biochemical quality parameters are also presented.

7.2 The Two-Way Error Component Regression Models

In the linear regression model $y_{ij} = \mu + x'_{ij}\beta + e_{ij}$, the error component e_{ij} may be decomposed as

$$e_{ij} = \alpha_i + \lambda_j + u_{ij} \ , \tag{7.1}$$

where α_i represents the effect due to i^{th} Taster and λ_j represents the effect due to j^{th} sample. Then the two-way error component regression model can be written as

$$y_{ij} = \mu + \alpha_i + \lambda_j + x'_{ij}\beta + u_{ij} \ , \tag{7.2}$$

$$i = 1, 2, \ldots, r \ , \ \ j = 1, 2, \ldots, n \ .$$

Here u_{ij} represents the remainder stochastic disturbance. We note here that α_i and λ_j are independent. In vector form, (7.2) can be written as

$$y = \mu 1_{nr} + Z_\alpha \alpha + Z_\lambda \lambda + \mathbf{X}\beta + u \ , \tag{7.3}$$

where Z_α and Z_λ are $(I_r \otimes 1_n)$ and $(1_r \otimes I_n)$ incidence matrices respectively. $\alpha' = (\alpha_1 \ \alpha_2 \ \ldots \ \alpha_r)$ *and* $\lambda' = (\lambda_1 \ \lambda_2 \ \ldots \ \lambda_n)$. We note here that $Z_\alpha \ Z'_\alpha = I_r \otimes J_n$ and $Z_\lambda \ Z'_\lambda = J_r \otimes I_n$. The regression coefficient vector β and the remainder disturbance u are as defined in the previous chapter.

7.2.1 The Fixed Effects Model

If α_i and λ_j are assumed to be fixed parameters and the remainder disturbance stochastic with $u_{ij} \sim iid(0, \sigma^2)$, then (7.2) represents a two-way fixed effects error component model. Here x_{ij} are assumed independent of u_{ij} for all i *and* j. Inference in this case would be specific to the n samples and r Tasters.

We note here that for large n *and* r, there will be too many dummy variables in the model, which may aggravate the multicollinearity problem among the regressors. Also,

there is an enormous loss of degrees of freedom as we have to estimate $n + r$ number of parameters apart from the k regression parameters. Baltagi (1995) discusses the fixed effects estimates of β by performing a transformation, which essentially sweeps the α_i and λ_j effects. The regression coefficient vector β and the remainder disturbance u are as defined in the last chapter.

7.2.2 The Random Effects Model

If we assume $\alpha_i \sim iid(0, \sigma_\alpha^2)$, $\lambda_j \sim iid(0, \sigma_\lambda^2)$ *and* $u_{ij} \sim iid(0, \sigma^2)$, independent of each other, then the model (7.2) becomes a two-way random effects regression model. Also, x_{ij} are independent of α_i, λ_j *and* u_{ij} for all i and j. In this case the inference is specific to the population from which the sample was drawn randomly. The variance-covariance matrix in this case would be

$$\begin{aligned} \Phi = E(e\, e') &= E(Z_\alpha \alpha + Z_\lambda \lambda + u)(Z_\alpha \alpha + Z_\lambda \lambda + u)' \\ &= Z_\alpha E(\alpha\, \alpha') Z_\alpha' + Z_\lambda E(\lambda\, \lambda') Z_\lambda' + u\, u' \\ &= \sigma_\alpha^2 (I_r \otimes J_n) + \sigma_\lambda^2 (J_r \otimes I_n) + \sigma^2 (I_r \otimes I_n)\ . \end{aligned}$$

Here the GLS estimator of β is BQUE, so also the estimators of variance components. Under the normality assumption, the ANOVA estimators of VCs are MVUE. Baltagi (1981a) performed a Monte Carlo study on a simple regression equation with two-way error component disturbances and studied the properties of OLS, the within estimator, MINQUE, and six feasible GLS estimators. According to the findings, the OLS estimator of β is unbiased, but asymptotically inefficient. The GLS estimator of β is BLUE, but the VCs can not be estimated simultaneously. All the feasible GLS estimators studied are asymptotically efficient. Paruch (1984) showed that the GLS estimator of β is asymptotically efficient as long as the estimate of σ^2 is consistent and the probability limits of the estimates of σ_α^2 and σ_λ^2 are finite.

7.3 The Two -Way Regression Model With Heteroscedastic Error

We have discussed so far the fixed and random effects models with homoscedastic error component(s). In this section we introduce the mixed effects formulation with heteroscedastic error. In the model (7.2) we suppose that the sample specific effects λ_js are fixed and needs to be estimated. The distributional assumptions may be made in two ways as we have done in the sixth chapter. We specify the two-way models on the basis of assumptions as follows :

Model A : $\alpha_i \sim iid\,(0, \sigma_\alpha^2)$, $u_{ij} \sim iid\,(0, \sigma_i^2)$ and $\sum_j \lambda_j = 0$

Model B : $\alpha_i \sim iid\,(0, \sigma_i^2)$, $u_{ij} \sim iid\,(0, \sigma_u^2)$ and $\sum_j \lambda_j = 0$.

The assumptions of independence of various components are same as before. Note that for both the models, the number of parameters to be estimated is $(n+k+r+2)$, as there are n sample specific effects, $k+1$ location parameters and $r+1$ variance components. The dispersion matrices for Model A and Model B would be same as those for Model I and Model II of the previous chapter respectively.

7.3.1 ML Estimation For Model A

As the assumptions about the random components in Model A are exactly same as those for Model I discussed in the previous chapter, the variance-covariance matrix for the mixed model

$$y = \mu 1_{nr} + Z_\alpha \alpha + Z_\lambda \lambda + \mathbf{X}\beta + u$$

would be

$$E(e\, e') = \Phi_A = diag(S_1\, S_2 \ldots S_r)\ ,$$

where S_i is a $n \times n$ matrix defined in Section 5.5.2 of Chapter 5. The determinant and inverse of S_i has already been presented in Chapter 5.

The loglikelihood function for Model A may be written as

$$l_A = -\frac{nr}{2} ln(2\pi) - \frac{n-1}{2}\sum_i ln\,\sigma_i^2 - \frac{1}{2}\sum_i ln(\sigma_i^2 + n\sigma_\alpha^2)$$

$$-\frac{1}{2}\sum_i\sum_j \frac{1}{\sigma_i^2}(y_{ij} - \mu - \lambda_j - x_j'\beta)^2 + \frac{1}{2}\sum_i \frac{1}{\sigma_i^2}\frac{\sigma_\alpha^2}{\sigma_i^2 + n\sigma_\alpha^2}[\sum_j (y_{ij} - \mu - \lambda_j - x_{ij}'\beta)]^2$$

$$= -\frac{nr}{2} ln(2\pi) - \frac{n-1}{2}\sum_i ln\sigma_i^2 - \frac{1}{2}\sum_i ln(\sigma_i^2 + n\sigma_\alpha^2) - \frac{n-1}{2}\sum_i \frac{1}{\sigma_i^2}S_i^2 +$$

$$\frac{n^2}{2}\sum_i \frac{1}{\sigma_i^2}\frac{\sigma_\alpha^2}{\sigma_i^2 + n\sigma_\alpha^2} D_i^2\,, \tag{7.4}$$

where $S_i^2 = \frac{1}{n-1}\sum_j (y_{ij} - \mu - \lambda_j - x_j'\beta)^2$, $D_i = \sum_j (y_{ij} - \mu - \lambda_j - x_j'\beta)$. We note here that the Ml estimates of the implicit parameters in this formulation can not be obtained directly as the model matrix is not of full rank. This follows from the following discussion.

Rank of Model Matrix ($\mathbf{X}$ Z_λ)

The augmented matrix of regressors and the incidence matrix associated with the fixed effects λ may be written as

$$(\mathbf{X}\ Z_\lambda) = \begin{bmatrix} X & I_n \\ X & I_n \\ \vdots & \vdots \\ X & I_n \end{bmatrix}\ ,\ \text{where } X \text{ is } (n \times k) \text{ matrix of regressors and } I_n \text{ is } (n \times n)$$

Now, $rank(\mathbf{X}\ Z_\lambda) = rank(X\ I_n)$, and the rank of this is atleast n because of last n columns. Again, as the maximum number of rows is n, the rank can not exceed n and hence the rank of the augmented matrix is n.

The Constrained ML Estimation

As we have to estimate $n+k+1$ number of fixed parameters, we need to impose $k+1$ restrictions to estimate the required number of fixed parameters.

We make some changes in the matrix of regressors for our convenience. Including the intercept term in the $\mathbf{X}$ matrix and the new matrix be denoted by $\mathbf{X}^*$ where the first

column is 1_{nr}. The model in this case may be written as

$$y = Z_\alpha \alpha + Z_\lambda \lambda + \mathbf{X}^* \beta^* + u \, .$$

Here the $\mathbf{X}^*$ matrix is of the form

$$\begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots & \\ \dots & \dots & \dots & \dots & \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{pmatrix} .$$

We now impose the restriction $\lambda' \mathbf{X}^* = \mathbf{0}$ which implies

$$\begin{aligned} \sum_j \lambda_j &= 0 \\ \lambda_1 x_{11} + \lambda_2 x_{21} + \dots + \lambda_n x_{n1} &= 0 \\ \lambda_1 x_{12} + \lambda_2 x_{22} + \dots + \lambda_n x_{n2} &= 0 \\ \dots \dots \dots \dots \dots \dots \dots & \\ \lambda_1 x_{1k} + \lambda_1 x_{2k} + \dots + \lambda_n x_{nk} &= 0 \end{aligned}$$

With the restriction the loglikelihood function may be defined as

$$l_A^* = l_A - \lambda' \mathbf{X}^* a \, ,$$

where $a = (a_0 \; a_1 \; \dots \; a_k)'$ is the vector of Lagrange Multiplier. Now,

$$l_A^* = -\tfrac{nr}{2} \, ln(2\pi) - \tfrac{(n-1)}{2} \textstyle\sum_i ln \, \sigma_i^2 - \tfrac{1}{2} \sum_i \, ln(\sigma_i^2 + n\sigma_\alpha^2)$$

$$-\tfrac{1}{2} \textstyle\sum_i \, \tfrac{1}{\sigma_i^2} \, \sum_j \, (y_{ij} - x_j^{*\prime} \beta^{*2} - \lambda_j)^2 + \tfrac{1}{2} \sum_i \, \tfrac{1}{\sigma_i^2} \, \tfrac{\sigma_\alpha^2}{\sigma_i^2 + n\sigma_\alpha^2} \, \left[\textstyle\sum_j (y_{ij} - x_j^{*\prime} \beta^* - \lambda_j) \right]^2 \qquad (7.5)$$

$$-a_1 \textstyle\sum_j \lambda_j \, x_{j1} - a_2 \sum_j \lambda_j \, x_{j2} - \dots - a_k \sum_j \lambda_j \, x_{jk}$$

Differentiating l_A^* with respect to λ_j and equating to zero and after simplification, we get

$$\lambda_j = (y^*_{jw} - x^{*\prime}_j\beta^*) - (\sum_i \frac{1}{\sigma_i^2})^{-1} \sum_i \frac{1}{\sigma_i^2} \frac{\sigma_\alpha^2}{\sigma_i^2 + n\sigma_\alpha^2} \sum_j (y_{ij} - x^{*\prime}_j\beta^*) \,, \qquad (7.6)$$

where $y^*_{jw} = \frac{\sum_i w_i y_{ij}}{\sum_i w_i}$ and $w_i = \sum_i \frac{1}{\sigma_i^2}$ is the weight. Here y^*_{jw} is the weighted mean for j^{th} sample taken over the r Tasters' scores and the weight is reciprocal to the error variances associated with individual Tasters' scores.

The ML estimators of σ_α^2 and σ_i^2 may be obtained from the following relations respectively :

$$\sum_i \frac{1}{\sigma_i^2 + n\sigma_\alpha^2} \left[1 - \frac{n}{\sigma_i^2 + n\sigma_\alpha^2} D_i^2\right] = 0 \text{ and}$$

$$\frac{\sigma_i^2}{\sigma_i^2 + n\sigma_\alpha^2} + \frac{\sigma_\alpha^2}{\sigma_i^2 + n\sigma_\alpha^2} \left[\frac{1}{\sigma_i^2} + \frac{1}{\sigma_i^2 + n\sigma_\alpha^2}\right] D_i^2 = \frac{1}{\sigma_i^2} \sum_j (y_{ij} - \lambda_j - x^{*\prime}_j\beta^*)^2 - (n-1).$$

7.4 Analysis of CTC Clonal Data

In this section we study the Data Set 9, which have been introduced in the first chapter. There are actually two data sets containing information on 30 Tocklai released CTC clones. A panel of three Tasters has evaluated the clones in terms of overall quality (V) and the scores are given on 0-10 point scale. The five biochemical parameters measured are TF, TR, C, CF and TLC. The plot of each biochemical parameters for the two data sets are presented in Fig 7.3 to Fig 7.7. The profile plot of Tasters' scores are presented in Fig 7.1 and Fig 7.2.

As may be observed from the graphical plots, the variation of TF values over the clones is very low, whereas that for TR is high. The caffeine and total liquor colour also varies over clones. However, the clonal variation of crude fiber is negligible. It is believed that the clones TV1, TV2 and TV17 are of good quality; TV 19 is of medium quality, and the rest are of average quality. This perception of the manufacturers is based on the price realization.

As initial diagnostics, we conduct the two-way ANOVA test to study if the clone wise variation is significant on the basis of Tasters' scores. The ANOVA results for both the data are presented in Table 7.1. Clearly the clonal variation is highly significant at 5% level. However, the average levels of Tasters' scores do not differ significantly as evident from the F-ratios for the Tasters. Thus, in associating the chemical information with the Tasters' choices, we may use the two-way error component regression model (Model A) which incorporates the sample wise variation as an assignable source of variation. For obvious reasons, we consider the effects due to clones as fixed.

We fit the Model A to both the data sets. The estimates of regression coefficients associated with different chemical parameters (along with t-ratios) and the estimates of variance components are presented in Table 7.2 for both the data sets. It may be observed that all the chemical quality parameters but crude fiber are statistically significant at 5% level in explaining the overall quality or value. This is true for both the samples collected in two consecutive years. It is interesting to observe that the coefficient estimates associated with CF are negative and statistically insignificant. This is against the common belief of the chemists, as they are of the opinion that in tea brew the CF is supposed to act positively towards quality and/or value and the Tasters are well in position to recognize the presence of CF. However, this notion of the chemists can not be questioned only on the basis of our analysis, as we are studying a single sample per clone. Had a large number of samples per clones been analyzed, then on the basis of regression analysis specific to a particular clone we could comment confidently on the behaviour of CF and other quality parameters. Also, the quality parameters measured for **green shoots** (e.g. total oxygen uptake, total carotenoids, etc.) needs to be analyzed separately to develop fairly good idea about the chemical characteristics of CTC clones and their association with the sensory choices.

The estimates of σ_α^2 are close to zero suggesting that the Tasters do agree on average about the quality. However, for the given data sets it means that the Tasters do not find any difference in quality for all the CTC clones. Anyway, the estimates of σ_i^2 are reasonably large and is highest for the first Taster (σ_1^2) for both sets of samples. The values of lnL and SSE suggests that the fit is better with the second set of samples. The SSE values are reasonably small in both the cases.

The estimated scores ($x_j^{*\prime}\hat{\beta}^* + \hat{\lambda}_j$) and the estimates of the clone specific fixed effects (λ_j) are presented in Table 7.3 and 7.4, along with the observed Tasters' scores. The measurements on biochemical quality parameters for both the data sets are presented in Table 7.5 and Table 7.6. Much technical discussions on the estimates of λ_j obtained can not be introduced at this instance, as we are no way in a position to detect those clones which appear to be the best in terms of quality, on the basis of given data sets. We can only say that the estimates of λ_j includes those clone specific effects with respect to several measured and unmeasured chemical components.

Table 7.1 : ANOVA Results for Two Sets of Data

	Source	d.f.	SS	F	$F_{critical}$
Data 1	Clone	29	86.08	6.30	1.66
	Taster	2	1.92	2.04	3.16
	Error	58	27.32		
Data 2	Clone	29	86.20	6.50	1.66
	Taster	2	0.68	0.75	3.16
	Error	58	26.50		

Table 7.2 : ML estimators of the regression coefficients and variance components

Parameters	Data 1		Data 2	
	Estimate	t-ratio	Estimate	t-ratio
$\widehat{\mu}^*$	11.9728	3.567	5.0473	2.978
$\widehat{\beta}_{TF}$	0.0411	2.67	0.8357	3.846
$\widehat{\beta}_{TR}$	0.2723	3.865	0.0031	1.968
$\widehat{\beta}_{C}$	0.3119	2.832	0.6539	2.573
$\widehat{\beta}_{CF}$	-0.3652	1.841	-0.2843	1.360
$\widehat{\beta}_{TLC}$	0.6615	3.102	0.6515	3.041
$\widehat{\sigma}_1^2$	1.1483		1.3331	
$\widehat{\sigma}_2^2$	0.5848		0.9438	
$\widehat{\sigma}_3^2$	0.9099		0.8786	
$\widehat{\sigma}_\alpha^2$	0.0001		0.0001	
$ln\ L$	-29.2082		-20.3158	
SSE	0.7283		0.4645	

Table 7.3 : The Tasters' scores along with the estimated scores and estimates of fixed effects λ_j for Data 1

Clone	Taster 1	Taster 2	Taster 3	Estimated Score	λ_j
CV1	9.30	8.10	8.00	8.3536	1.0998
CV2	8.70	7.50	7.00	7.6341	1.1971
CV3	6.50	6.00	6.50	6.2671	-0.1521
CV4	6.00	6.50	6.00	6.2318	-0.0447
CV5	7.50	7.00	6.50	6.9685	0.2497
CV6	7.00	7.00	6.00	6.7008	0.4054
CV7	7.20	6.50	6.00	6.5158	-0.7544
CV8	7.50	6.00	5.50	6.2051	-1.0721
CV9	9.10	8.00	8.37	8.3702	0.9753
CV10	7.00	7.50	6.00	6.9331	0.3058
CV11	7.80	7.50	7.00	7.4211	0.1509
CV12	7.60	6.50	7.00	6.9091	0.0019
CV13	6.00	6.50	8.00	6.8291	0.6305
CV14	6.50	6.00	7.00	6.4164	-0.4929
CV15	7.00	8.00	6.50	7.3148	-0.4509
CV16	7.20	7.00	5.00	6.4495	-0.8821
CV17	7.00	7.50	7.00	7.2318	-0.0631
CV18	9.20	7.20	8.40	8.0311	1.5861
CV19	6.00	6.50	6.00	6.2318	-0.1161
CV20	8.50	8.00	7.50	7.9684	0.6823
CV21	6.50	7.00	7.20	6.9409	0.4534
CV22	7.50	6.50	7.10	6.9153	0.3153
CV23	7.00	6.30	6.00	6.3755	-0.0894
CV24	6.20	5.00	5.50	5.4327	-0.6817
CV25	5.50	5.00	6.50	5.5658	-0.1886
CV26	5.00	4.00	5.00	4.5348	-1.5438
CV27	4.00	6.50	5.50	5.6091	-0.2169
CV28	4.00	6.00	5.00	5.2275	0.3866
CV29	5.00	5.60	5.50	5.4276	-1.0301
CV30	5.00	6.00	5.00	5.4642	-0.7904

Table 7.4 : The Tasters' scores along with the estimated scores and estimates of fixed effects λ_j for Data 2

Clone	Taster 1	Taster 2	Taster 3	Estimated Score	λ_j
CV1	8.60	7.90	9.15	8.5606	1.9448
CV2	8.90	8.50	7.50	8.2156	1.0817
CV3	7.20	7.10	6.50	6.8937	-0.1157
CV4	6.50	5.90	7.20	6.5545	-0.5023
CV5	6.90	7.20	6.20	6.7375	-0.5000
CV6	7.20	6.80	5.90	6.5542	-0.4749
CV7	6.20	6.80	6.50	6.5314	-..1246
CV8	5.60	6.00	5.80	5.8209	-0.6024
CV9	8.70	8.50	6.90	7.9330	1.2053
CV10	6.50	6.30	7.20	6.6983	0.3224
CV11	5.20	4.50	4.90	4.8325	-1.1823
CV12	5.90	5.80	4.50	5.3234	-0.9831
CV13	5.88	6.25	6.55	6.2716	0.4578
CV14	5.80	5.25	6.25	5.7760	-0.6902
CV15	4.50	4.20	4.80	4.5079	-1.4968
CV16	5.20	5.00	5.90	5.3983	-0.6048
CV17	6.50	6.00	6.80	6.4361	0.4516
CV18	8.90	8.50	7.80	8.3314	2.2494
CV19	6.85	6.25	5.88	6.2598	-0.3683
CV20	5.70	5.70	7.50	6.3949	-0.2896
CV21	6.50	6.00	5.50	5.9341	-0.8074
CV22	6.50	7.30	7.30	7.0964	0.9872
CV23	8.50	6.30	6.50	6.9369	0.7042
CV24	6.20	6.80	5.80	6.2612	-0.2585
CV25	6.70	6.00	5.80	6.1008	-0.4397
CV26	8.00	6.50	6.20	6.7658	0.4775
CV27	7.60	6.30	6.00	6.5149	0.02139
CV28	4.00	4.90	5.60	4.9421	-0.3323
CV29	4.55	5.89	6.50	5.7844	-0.3858
CV30	5.50	7.50	5.80	6.3346	0.2370

Table 7.5: The biochemical measurements for Data 1

CLONE	TF	TR	CAF	CF	TLC
TV1	1.77	14.03	3.52	9.54	6.50
TV2	1.88	12.23	3.65	10.40	6.40
TV3	1.11	11.38	3.92	10.70	6.62
TV4	1.05	11.60	3.74	10.40	6.47
TV5	1.78	12.84	3.45	10.64	6.23
TV6	1.54	12.34	3.75	10.81	6.55
TV7	1.55	14.08	3.60	10.35	6.05
TV8	1.70	12.95	3.59	8.98	6.34
TV9	1.53	12.86	3.82	8.60	6.32
TV10	1.43	11.52	3.48	8.96	6.73
TV11	1.65	13.96	2.99	9.06	6.73
TV12	1.58	12.15	3.69	9.09	6.50
TV13	1.30	11.13	3.56	10.46	6.38
TV14	1.59	14.09	3.50	10.23	6.67
TV15	1.49	14.62	3.27	9.49	6.00
TV16	1.59	13.43	3.19	9.57	6.13
TV17	1.50	12.48	3.43	9.00	6.29
TV18	1.27	10.95	3.68	9.68	6.36
TV19	1.44	10.70	3.97	9.57	6.47
TV20	2.06	12.48	4.00	8.76	6.27
TV21	1.96	11.60	4.06	9.86	6.50
TV22	1.74	12.66	3.92	9.67	6.86
TV23	1.70	12.26	3.97	9.50	6.99
TV24	1.45	12.78	3.71	10.90	6.95
TV25	1.27	11.48	3.87	11.00	6.89
TV26	1.32	12.38	3.79	10.23	7.20
TV27	1.05	11.84	3.85	11.03	6.90
TV28	0.88	8.96	3.62	11.85	6.72
TV29	1.24	8.80	3.87	8.92	5.87
TV30	1.12	8.59	4.17	8.88	6.10

Table 7.6: The biochemical measurements for Data 2

CLONE	TF	TR	CAF	CF	TLC
TV1	1.46	14.67	3.61	12.06	5.85
TV2	1.57	14.46	4.71	12.58	5.90
TV3	1.11	11.62	4.61	11.84	4.92
TV4	1.20	12.05	4.57	11.94	5.19
TV5	1.41	14.63	4.52	11.54	5.60
TV6	1.44	13.55	3.85	10.93	5.73
TV7	1.16	12.32	3.76	10.92	4.89
TV8	1.21	11.60	4.14	13.32	4.83
TV9	1.28	12.25	4.16	11.28	4.91
TV10	1.18	11.12	4.03	11.24	4.36
TV11	1.26	11.78	3.59	11.82	4.60
TV12	1.32	11.57	4.31	12.34	4.63
TV13	0.98	9.98	4.08	12.37	3.69
TV14	1.27	12.27	4.41	11.59	4.39
TV15	1.19	12.70	3.39	11.13	4.39
TV16	1.31	13.41	3.37	11.18	4.56
TV17	1.26	13.07	3.72	11.69	4.36
TV18	0.98	09.77	4.09	11.06	3.52
TV19	1.25	11.40	4.30	10.84	4.39
TV20	1.70	14.55	3.56	11.18	5.93
TV21	1.64	12.68	3.71	11.32	5.86
TV22	1.47	12.62	3.18	11.06	5.09
TV23	1.47	12.45	3.74	11.43	4.88
TV24	1.12	12.79	4.36	11.55	4.30
TV25	1.25	11.95	4.39	11.59	4.49
TV26	1.15	12.52	3.65	11.74	4.78
TV27	0.98	10.91	4.21	11.94	4.41
TV28	0.78	8.80	3.63	12.30	3.03
TV29	0.91	10.87	3.35	12.29	4.48
TV30	0.82	9.86	3.45	11.70	4.26

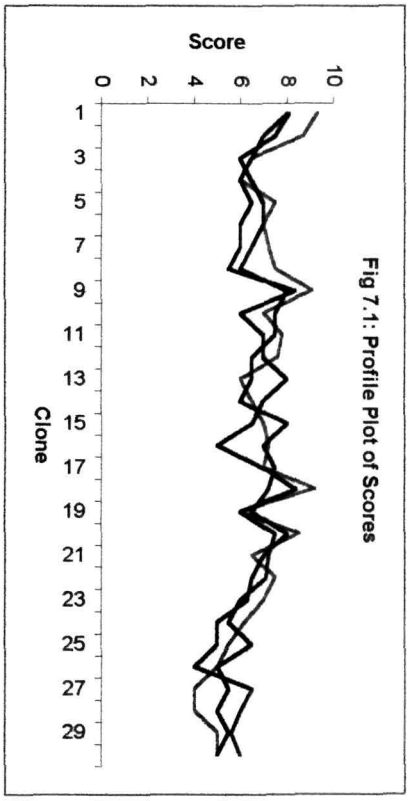


Fig 7.1: Profile Plot of Scores

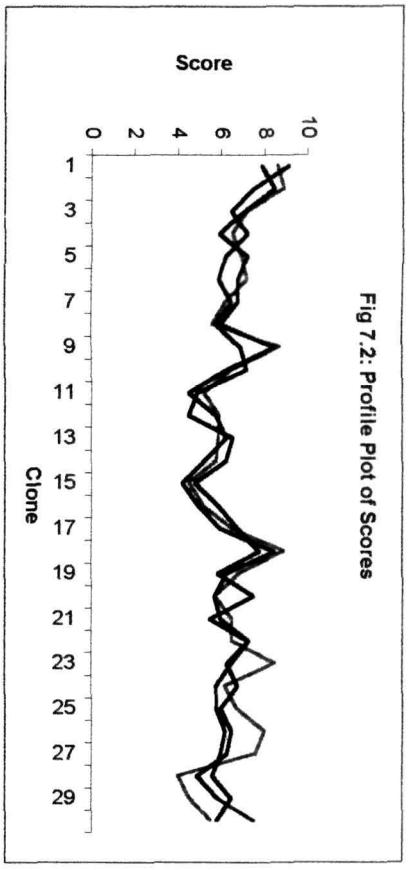


Fig 7.2: Profile Plot of Scores

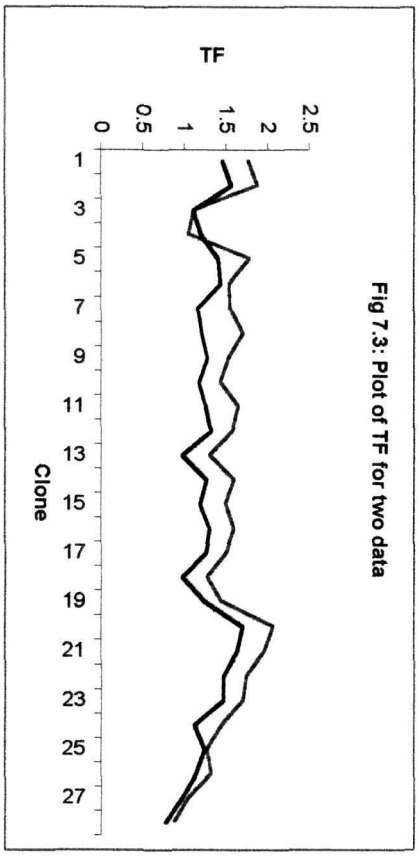


Fig 7.3: Plot of TF for two data

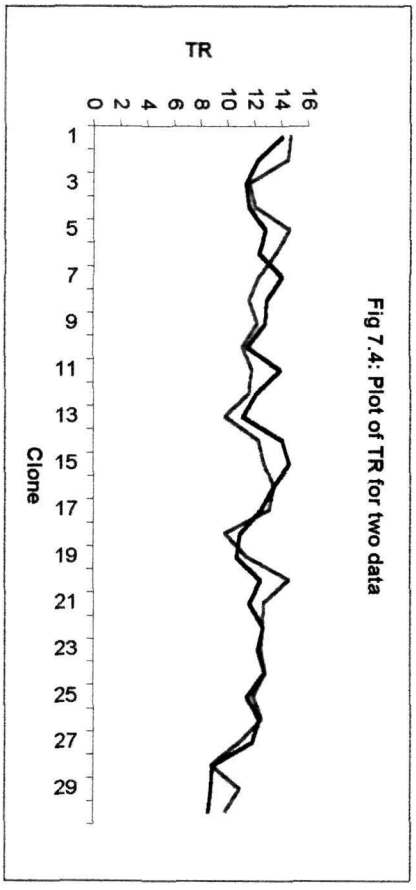


Fig 7.4: Plot of TR for two data

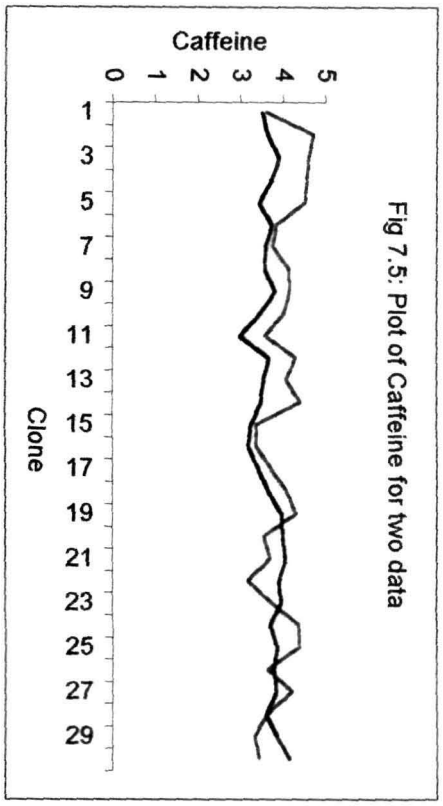


Fig 7.5: Plot of Caffeine for two data

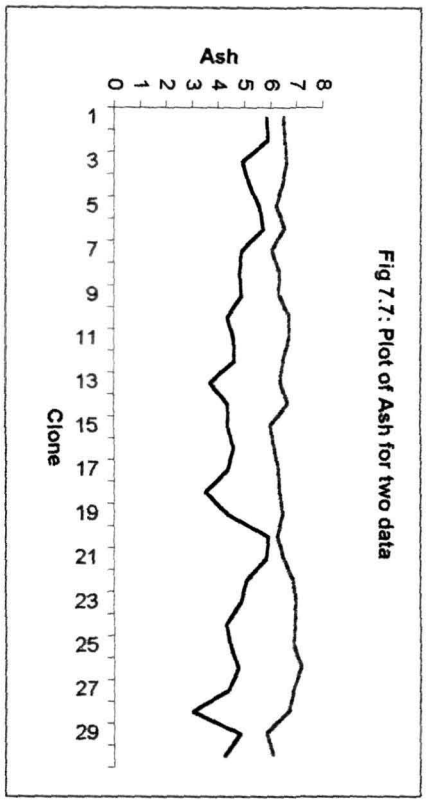


Fig 7.7: Plot of Ash for two data

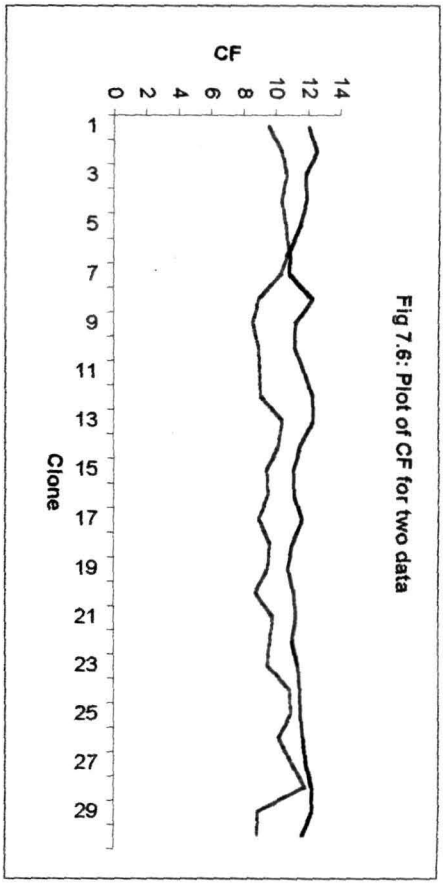


Fig 7.6: Plot of CF for two data

CHAPTER - 8

CONCLUDING DISCUSSIONS

In this study we have tried to understand the association between the chemistry in tea and the sensory choices. This is typically a quality determination problem in tea industry, which has not received much attention from the scientific community. We address this problem and try to solve this practical problem using statistical methodologies. Our main aim was to understand if the effects of the important biochemical quality parameters (according to the belief of the chemists) are reflected, may be partially, in the Tasters' choice. This is because, if the significance of chemical parameters in ultimate quality assessment can be established, then we may expect that the whole system of quality determination and the auction pricing can be rationalized to some extent. If some threshold limits for few important parameters can be evolved, which may guide the chemists as well as the sensory experts to understand the possible levels of different quality attributes for the given set of samples, then the absolute dependence on sensory analysis can be minimized. In fact, these threshold limits, however approximate they may be, may guide the brokers and the buyers about the quality of a given lot of tea.

To start with, we have tried to understand the association of few important chemical parameters with the Taster's scores using the statistical regression techniques. Approximate linear relation between the chemical measurements and the sensory scores on various quality attributes have been observed in almost all the data sets studied. One may raise question on such associations on the ground that many other chemical parameters responsible for aroma and strength in CTC teas, have not been considered. Such questions are quite justified. Apart from other phenolic compounds, different levels of theaflavins and thearubigins also needs to be included in the regression formulation to get a better understanding of the association. However, we had to depend on the data provided by the Tea Research Association of India, and the insufficiency of chemical information restricts our study and general conclusion regarding the behaviour of chemical parameters

can not be drawn. But, we have observed that the total levels of TF and TR significantly contribute towards quality. One point worth mentioning here that the Caffeine aspect, as much talked about in case of CTC teas, needs to be studied in detail in relation to different levels of TF and TR. It is known that in the tea brew, part of the caffeine must complex with flavanols (TF, TR, etc.) and play an important role in the tea taste, with contributions to briskness, mouthfeel, and thickness. As discussed by Yamanishi (1995), TF and TR are very astringent, and caffeine is very bitter. Interaction between these compounds reduce the astringency of the former and the bitterness of the latter. It is claimed by the chemists that the caffeine contributes about one fourth of the bitterness of the tea brew. Removal of caffeine from the tea infusion has a significant effect on the taste of the infusion. Decaffeination cause the bitterness of a black tea infusion to increase slightly and changes the nature of the astringency.

It is necessary to measure different TF and TR levels along with Caffeine and the pattern of their variation along with the possible interactions needs to be understood thoroughly. A fairly good amount of information on such behaviour may be obtained from the graphical plots. It is possible to study the nature of interactions and their effects on different quality attributes using regression techniques. Given different samples collected from various gardens, the Pattern Recognition Technique may be also applied to understand the variation in chemical behaviour. A good idea about the techniques used to understand the variations in chemical measurements may be obtained from the web site www.chemometrics.com. Understanding the interrelationship among various chemical parameters is very important before we associate chemistry with the sensory analysis. We have only tested the multicollinearity among few chemical parameters available with us. The influence of various volatile and non-volatile compounds and their possible interactions could not be studied, which obviously limits our study. It is important to see the peaks of various TF and TR in HPLC chromatogram while assessing the variations. However, we could not obtain such information.

As the sensory choices are subjective, to understand the sample wise variation in different quality attributes we need a sensory panel data. A group of Tasters' choices is needed to statistically assess the between-Taster and within-Taster variability. Replication of choices on each sample by each Taster would be of much help to track different possible aspects of variations. Had replicated scores been available, we could go for a more detailed discussions about possible statistical formulations to address the problem of reproducibility. Also, to develop a proper statistical methodology to measure the Tasters' precision, we need replicated scores. Unless, all possible variations in sensory panel data are eliminated, specific to a particular quality attribute, the application of variance components models would not be of much help. Also, the use of error-component regression models (while simultaneously studying the chemical parameters and the sensory choices) would not provide sufficient information about the significance of different chemical parameters unless the quality assessment experiment is properly designed taking care of all possible aspects of product variation and the variation due to Tasters.

In the past 20 years there has been a considerable increase in the activity in the field of sensory evaluation. Evidence for this is seen by the number of books, journals, and articles published on the topic; the number of professional organizations and the number of universities offering sensory courses. Much of the recent growth for sensory evaluation can be attributed to the increased interest and support of the consumer product industry. The food and beverage industry provides the vital sponsorship for the sensory activities. For this industry, sensory analysis is the natural extension of each company's aim to achieve highest product quality and thus attain a dominant share in the highly competitive market. The rapidly developing technology and the subtleties of the market dictate that all available resources, including sensory evaluation, be used to the best advantage.

We now focus on certain problems of measuring consumer acceptance for a specific brand of tea (or a food product in general). A tea producing company may be interested to know whether the consumers can differentiate their product with the other available varieties. Such knowledge is very important from the **marketing** point of view. In

this case, the ideal practice would be to design an experiment to assess the different quality aspects of the particular brand of tea, where the choices of a large number of consumers along with the choices of the sensory panelists needs to be analyzed. As the consumers' way of preparing tea differs from the Tea Tasters' method of tea preparation, the comparison of the choices given on two completely different preparations would not help. Here the treatment differs. That is why the sensory panelists should also taste the **home-type prepared** teas. The company's aim would always be to study whether the consumers can identify the inherent quality or qualities of the brand under consideration in the same way the sensory panelists can. This is purely a choice identification problem which can be solved statistically if the information on the choices are obtained properly on the basis of a suitably designed experiment.

Coinciding with these activities, marketing strategies need to be evolved. This is certainly related to the company's sales and advertising activities. On the basis of the findings of choice determination experiments, we may always obtain a fairly good idea about the consumers' choices. Thus the advertising policy would be based on those preferences. The product wise and region wise choice variation may always be taken care of as the consumer segments and niche marketing are important concepts in marketing research. The statistical techniques may provide great support to the marketing specialists in order to better comprehend the consumer behaviour.

Thus the importance of sensory analysis can not be denied. But a proactive and less fragmented approach to applying sensory analysis will require testing large number of products, as mentioned earlier, to provide more comprehensive learning about how product and process variables influence the consumers' perception and preferences. The statistical models will certainly describe the relationship between perception and product variables with greater understanding and more long term benefits is expected for the product development process.

We now turn our discussion to an important inference problem in the error component regression models discussed in this study. The one-way random effects model is used when there is sufficient ground to believe that the tea samples are truly random sam-

ples representing a particular population, and the variation due to sample is negligible. However, for a given data we may always try both the one-way and two-way models and compare the fits. The comparison of SSE values and the likelihood based AIC criteria seem sufficient in this situation. One important aspect which needs to be addressed here is the **consequence of under-specifying or over-specifying the error component model on the variance components estimates.**

Underspecifcation : Here the true model is two-way with

$$e_{ij} = \alpha_i + \lambda_j + u_{ij}$$

$$i = 1, 2, \ldots r \, , j = 1, 2, \ldots n$$

while the estimated model is one-way with random component

$$e_{ij} = \alpha_i + u_{ij} \ ,$$

$\sum \lambda_j = 0$, $\alpha_i \sim iid(0, \sigma_\alpha^2)$, $u_{ij} \sim iid(0, \sigma_i^2)$, independent of each other among themselves.

As may be observed from Chapter 7, the ML estimators of σ_i^2 involves λ_js, the fixed parameters. Here one interesting study would be to estimate the bias of σ_i^2 *and* σ_α^2 for the misspecified model. The consistency and unbiasedness of the variance components under different estimation methods (ML, two-stage least squares, etc.) may also be studied. Similarly we may study the behaviour of variance components under over-specification. Also, the small sample behaviour of these estimators needs to be addressed thoroughly. However, we could not address all these theoretical problems in this thesis, and there is much scope for further study.

Another important aspect is the **Determination of superior CTC clones.** In Chapter 7 we have studied the quality aspects in Tocklai released CTC clones. The manufacturers believe that among these clones, only a few (CV 1, CV 2 and CV 17) yield best in terms of quality and market value. Their observation is simply based on the price realization. The Tocklai Experimental Station has been studying the chemical behaviour of these clones on a continuous basis. The general practice is to measure the chemical

parameters for a single sample for each clone. The Taster(s) also give single score on each sample. A typical presentation of this practice is the data we have studied in the seventh chapter.

If interest lies in determining the best clones in terms of market performance, then we believe that a more comprehensive approach is necessary. It is possible to determine the best performing clones in terms of quality or overall quality using the advanced statistical techniques like Frontier Techniques and Data Envelopment Analysis (DEA). The frontier technique is related to the econometric concepts of frontier production functions and technical efficiency measurements. In DEA the basic approach is non-parametric, where there is a mapping of chemical information on sensory choices without any prior assumption on the distributions of errors and on specific functional relationship.

Suppose for each of the 30 clones, several samples are studied in terms of chemical characteristics and sensory choices. All the samples under study must be evaluated by the sensory panel. If such data base is available, we may formulate our problem as - given the measured levels of chemical parameters, what would be the possible maximum score in terms of quality/market value. In doing so we may always take care of the bias associated with individual members of sensory panel. Such studies would certainly help the manufacturers to rationalize the system of "best product determination" which would ultimately help them in introducing their products in market with confidence. Also, the introduction of a methodological approach in product selection would minimize the dependence on sensory panelists. This practical problem can be solved using the above mentioned statistical techniques.

We have tried to understand the quality aspects in CTC tea in this study with the data base provided by the Tea Research Association of India. We do not claim any originality in our statistical formulations. We have only extended the developed techniques in some situations to fit our empirical set up. We should say that statistics as a tool has been applied with confidence in understanding the chemical and sensory information related to tea quality assessment.

Note 1

One interesting statistical formulation related to non-linear model formulation is the **Box-Cox model with error associated with the response**. In Chapter 4, we have mentioned this while associating the chemical components in tea with the single Taster's sensory choice. Here we try to formulate a transformation model introducing error component with the response variable. This is done keeping in view the subjectivity associated with the Taster's choice. Only transformation of the response variable with additive error component is considered. We note here that this formulation is not presented in the Chapter 4 only because of the fact that a local convergence for the estimates of the implicit parameters by maximizing the loglikelihood function could not be achieved.

Suppose the true values of the response are $Y_i (i = 1, 2, \ldots n)$, which are unknown. The observed response are y_i. We assume $y_i = Y_i + v_i$, where v_i's are the errors-in-variables. We suppose a functional relationship between the regressors and the response variable as

$$\frac{Y_i^{\lambda} - 1}{\lambda} = \beta_0 + \beta_1 \, x_{1i} + \ldots + \beta_k \, x_{ki} + e_i, \tag{1}$$

where λ is the Box-Cox type transformation parameter. Assuming $e_i \sim N(0, \sigma_e^2)$, σ_e^2 being unknown, we have

$$g(e_i) = \frac{1}{\sqrt{2\pi}\,\sigma_e} \, exp\,[-e_i^2/2\sigma_e^2].$$

Now from (1) we have

$$e_i = \frac{Y_i^{\lambda} - 1}{\lambda} - \beta_0 - \beta_1 \, x_{1i} \ldots \beta_k \, x_{ki},$$

which give $\frac{\partial e_i}{\partial Y_i} = Y_i^{\lambda - 1}$. Thus we may write the probability differential of Y_i as

$$f(Y_i) = \frac{1}{\sqrt{2\pi}\,\sigma_e} \, exp \left[-\frac{1}{2\sigma_e^2} \left\{ \frac{Y_i^{\lambda} - 1}{\lambda} - \beta_0 \ldots - \beta_k \, x_{ki} \right\}^2 \right] \, Y_i^{\lambda - 1} \, dY_i \tag{2}$$

We now consider the conditional distribution of the error component v_i given each Y_i to be Pearsonian type II. Then the conditional density function may be written as

$$p(v_i/Y_i) = Y_i^{-1}\, K_m \left[1 - \left(\frac{v_i}{Y_i}\right)^2\right]^m dv_i. \tag{3}$$

where,

$$K_m = B^{-1}(m+1, \frac{1}{2}) = B^{-1}(m+1, m+1)\; 2^{-(2m+1)} \tag{4}$$

is a beta function.

The joint density function of v_i and Y_i may now be written as

$$\begin{aligned} f(v_i\,,\, Y_i) &= p(v_i/Y_i).g(Y_i) \\ &= Y_i^{-1}\, K_m \left[1 - \left(\tfrac{v_i}{Y_i}\right)^2\right]^m \tfrac{1}{\sqrt{2\pi}\sigma_e} \\ exp\left[-\tfrac{1}{2\sigma_e^2}\left\{\tfrac{Y_i^{\lambda}-1}{\lambda} - \beta_0 - \beta_1\, x_{1i} \ldots - \beta_k\, x_{ki}\right\}^2\right] & Y_i^{\lambda-1}. \end{aligned} \tag{5}$$

Now to find the Jacobian of transformation, we proceed as follows :

$$\begin{aligned} y_i &= Y_i + v_i \\ \Rightarrow Y_i &= y_i - v_i = y_i - \frac{v_i}{Y_i}\, Y_i = y_i - w_i Y_i \end{aligned}$$

$$\frac{\partial y_i}{\partial Y_i} = 1, \frac{\partial y_i}{\partial v_i} = 1, \frac{\partial w_i}{\partial Y_i} = -\frac{v_i}{Y_i^2}, \frac{\partial w_i}{\partial v_i} = -\frac{1}{Y_i}.$$

The Jacobian of transformation is

$$\frac{\partial(y,\, w)}{\partial(Y,\, v)} = \begin{vmatrix} 1 & 1 \\ -\frac{v_i}{Y_i^2} & \frac{1}{Y_i} \end{vmatrix} = \frac{1}{Y_i}(1 + w_i)$$

The joint density function of w_i and y_i can now be written as

$$
\begin{aligned}
f(w_i,\, y_i) &= Y_i^{-1}\, K_m [1 - w_i^2]^m \;\; Y_i^{\lambda-1} \left(\tfrac{Y_i}{1+w_i}\right) \tfrac{1}{\sqrt{2\pi}\,\sigma_e} \\
&\quad exp\left[-\tfrac{1}{2\sigma_e^2} \left\{ \frac{\left(\frac{y_i}{1+w_i}\right)^{\lambda} - 1}{\lambda} - \beta_0 - \beta_1\, x_{1i} \ldots - \beta_k\, x_{ki} \right\}^2 \right] \\
&= \tfrac{1+w_i}{y_i} \left(\tfrac{y_i}{1+w_i}\right)^{\lambda} \tfrac{1+w_i}{y_i} (1 - w_i^2)^m\, K_m \tfrac{y_i}{(1+w_i)^2} \tfrac{1}{\sqrt{2\pi}\,\sigma_e} \\
&\quad Exp\left[-\tfrac{1}{2\sigma_e^2} \left\{ \frac{\left(\frac{y_i}{1+w_i}\right)^{\lambda} - 1}{\lambda} - \beta_0 - \beta_1\, x_{1i} \ldots - \beta_k\, x_{ki} \right\}^2 \right] \qquad (6) \\
&= \tfrac{1}{y_i} \left(\tfrac{y_i}{1+w_i}\right)^{\lambda} K_m\, (1 - w_i^2)^m \tfrac{1}{\sqrt{2\pi}\sigma_e} \\
&\quad exp\left[-\tfrac{1}{2\sigma_e^2} \left\{ \frac{\left(\frac{y_i}{1+w_i}\right)^{\lambda} - 1}{\lambda} - \beta_0 \ldots - \beta_k\, x_{ki} \right\}^2 \right]
\end{aligned}
$$

Now the density function y_i may be obtained by integrating the joint density function $f(w_i, y_i)$ with respect to w_i as

$$p(y_i) = \int_R f(w_i,\, y_i)\, dw_i.$$

The likelihood function is

$$
\begin{aligned}
L &= \textstyle\prod_{i=1}^n p(w_i) \\
&= \textstyle\prod_{i=1}^n \int_R \tfrac{1}{y_i} \left(\tfrac{y_i}{1+w_i}\right)^{\lambda} K_m (1 - w_i^2)^m \tfrac{1}{\sqrt{2\pi}\sigma_e} \qquad (7) \\
&\quad exp\left[-\tfrac{1}{2\sigma_e^2} \left\{ \tfrac{(y_i 1+w_i)^{\lambda} - 1}{\lambda} - \beta_0 - \beta_1\, x_{1i} \ldots \beta_k\, x_{ki} \right\}^2 \right] dw_i.
\end{aligned}
$$

We have tried to maximize the log of the likelihood function (7) to obtain the estimates of the parameters m, λ, σ_e^2, *and* $k+1$ β coefficients using the SPLUS computational package. However, for the given data set (studied in Chapter 4), the local convergence could not be achieved.

REFERENCES

1. Airy G B (1861): On the algebraical and numerical theory of errors of observations and the combinations of observations. McMillian, London.

2. Aiken L S and West S G (1991): Multiple Regression Testing and Interpreting Interactions. *Newbury Park*, Sage Publications.

3. Amerine M A and Pangborn R M (1965): Principles of Sensory Evaluation of Food. *Academic Press*, New York.

4. Anderson R D (1978): Studies on the estimation of variance components. Ph. D Thesis, Carnell University, New York.

5. Anderson R L and Bancraft T A (1952): Statistical Theory in Research. McGraw-Hill, New York.

6. Anderson T W (1982): An introduction to multivariate analysis. John Wiley and Sons, New York.

7. Arnold S F (1981): The Theory of Linear Models and Multivariate Statistics. John Wiley and Sons, New York.

8. Azzalini A (1984): Estimation and hypothesis testing for collections of autoregressive time series. *Biometrica*, 71, 85-90.

9. Avery R B (1977): Error component and seemingly unrelated regression. *Econometrica*, 45, 47-57.

10. Akike H (1973): Information theory and the extension of ML principle in B.N. Petrov and F. Csaki, eds. , 2nd International Conference on Information Theory, Akailseoniaikiuds, Budapest, pp. 267-281.

11. Baruah D N (1992): Tea Quality. *Two And A Bud*, 39(2), 2-6.

12. Baltagi B H (1982a): Prediction with two-way error component regression model. *Econometric Theory*, 4, 177.

13. Baltagi B H and Griffen J M (1988): A generalized error component model with heteroscedastic disturbances. *International Economic Review*, 29, 745-753.

14. Baltagi B H (1995): Econometric Analysis of Panel Data. John Wiley and Sons.

15. Bartlett M S (1937): Properties of sufficiency and statistical tests. *Proceedings of the Royal Society*, Series A, 160, 268-282.

16. Balsera P (1973): Best quadratic unbiased estimators of the variance-covariance matrix in normal regression. *Journal of Econometrics*, 2, 17-28.

17. Bates D M and watts D G (1981): A relative offset orthogonality convergence criterion for nonlinear least square. *Technometrics*, 23, 179-183.

18. Beilter P J and Landis J R (1985): A mixed effects model for categorical data. *Biometrics*, 41, 991-1000.

19. Biswas A K and Biswas A K (1971): Statistical evaluation of biological constituents and their effects on colour, brightness and strength. *J Sci. Food Agric.*, 24, 1457-1477.

20. Brockhoff P M and Skovgaard Ib M (1994) : Modelling individual differences between assessors in sensory analysis. *Food Quality And Preference*, 5, 215-224.

21. Breusch T S and Pagan A R (1980): The LM test and its application to model specification in econometrics. *Review of Econometric Studies*, 47, 239-253.

22. Brown P A W, Kleidon T A and Marsh A W (1983): New evidence on the nature of size related anamolies in stock prices. *Journal of Finantial Economics*, 12, 38-56.

23. Brown P A W, and Forsythe A B (1974) : Robust tests for the equality of variances. J A S A, 69, 364-367.

24. Box G E P and Cox D R (1964): An analysis of transformation. *J. Royal Statist. Soc.*, Series B, 35, 473-479.

25. ——— (1982): An analysis of transformation revisited, rebutted. *J A S A*, 77, 209-210.

26. Buning H (1997): Robust analysis of variance. *Journal of Applied Statistics*, 24(3), 319-332.

27. Bishop T A and Dudewicz E J (1978): Exact analysis of variance with unequal variances. *Technometrics*, 20, 419-424.

28. Carrol R J (1982): Tests for regression parameters in power transformation model. *Scan. J. Statist.*, 9, 217-222.

29. ——— (1980): A robust method of testing transformations to achieve approximate normality. *J. Royal Statist Soc.*, Series B, 42, 71-78.

30. Carrol R J, Ruppert D and Stefanski L A (1995): *Measurement Error in Non-linear Models.* Chapman and Hall, London.

31. Carrol R J and Ruppert D (1981): On prediction and power transformation family. *Biometrica*, 68, 609-615.

32. ——— (1984): Power transformation when fitting theoretical model to data. *J A S A*, 79, 312-328.

33. ——— (1987): Diagnostics and robust estimation when transforming the regression model and response. *Technometrics*, 3, 287-300.

34. ——— (1982): A comparison between maximum likelihood and generalized least squares in heteroscedastic linear model. *J A S A*, 77, 878-882.

35. Chauvert W (1863): A Manual of Spherical and Practical Astronomy. *LippinCott, Philadelphia.*

36. Chaudhary H, Borah M and Paul S K (1999): Analysis of tea quality data using multiple response regression model. To appear in the *Research Journal of A S S.*

37. Cochran W G (1939): The use of analysis of variance in enumerating by sampling. *J A S A*, 34, 492-510.

38. Crump S L (1947): The Estimation of Variance in Multiple Classification. Ph. D. Thesis, Iowa State University, Iowa.

39. —— (1951): The present status of variance components analysis. *Biometrics*, 7, 1-16.

40. Crowder M J and Hand D J (1990): Analysis of Repeated Measures. Chapman and Hall, London.

41. Chen G and Lockhart R A (1997): Box-Cox transformed linear models : A parameter based asymptotic approach. *Canadian Journal of Statistics*, 25(4), 517-529.

42. Collins S (1991): Prediction technique for Box-Cox regression models. *Journal of Business and Economic Statistics*, 9, 267-277.

43. Conniffe D (1982a): Testing the assumptions of seemingly unrelated regression. *Review of Economic Statistics*, 64, 172-174.

44. Cox D R and Reid N (1987): Parameter orthogonality and conditional inference. *J. Royal Statist. Soc.*, Series B, 49, 302-309.

45. Cole J W and Grizzle J E (1966): Application of multivariate ANOVA to reapeted measures experiments. *Biometrics*, 22, 810-828.

46. Corbeil R R and Searle S R (1976b): A comparison of variance components estimators. *Biometrics*, 32, 779-791.

47. Cloughley J B (1981): Storage deterioration in Central African tea: changes in chemical composition, sensory characteristics and price evaluation. *J. Sci. Food Agric.*, **32**, 1213-1223.

48. Cullis B R and McGilchrist C A (1990): A model for the analysis of growth data from designed experiment. *Biometrics*, 46, 131-142.

49. Chow G C (1960): Test of equality between two sets of coefficients in two linear regression. *Econometrica*, 28(3), 591-605.

50. Das A K, Gogoi R and Goswami M R (1994): Modification of CTC processing for quality improvement. *Proceedings of 32 nd Tocklai Annual Conference*,394-399.

51. Davis A G (1983): Theaflavins, objective indicators. *Tea and Coffee Trade Journal*, 155, 34.

52. Daniels H E (1939): The estimation of components of variance. *J. Royal Statist. Soc.*, Suppl. 6, 187-197.

53. Deb S B and Ullah M R (1986): The role of TF and TR in the evaluation of black teas. *Two And A Bud*, **15**, 101-103.

54. Dempster A P, Laird N M and Rubin D B (1977): Maximum likelihood from incomplete data via the EM algorithm. *J. Royal Statist. Soc.*, Series B, 39, 1-38.

55. Dempster A P, Rubin D B and Tsukawa R K (1981): Estimation of covariance components models. *J A S A*, 76, 341-353.

56. Dempster A P, Selwyn M R and Patel C M (1984): Statistics and computational aspects in mixed model analysis. *Applied Statistics*, 33, 203-214.

57. Dielman T E (1989): Pooled Cross-Section and Time Series Data Analysis. *Marcel Dekker Inc*, New York.

58. Diggle P J, Liang K Y and Zeger S L (1995): Analysis of Longitudinal Data. Clarendon Press, Oxford.

59. Diggle P J (1988): An approach to the analysis of repeated measurements. *Biometrics*, 44, 959-971.

60. Ellis R T and Cloughley J B (1981): The importance of theaflavins in tea liquors. *International Tea Journal*, 2, 7-8.

61. Fearn T (1975): A bayesian approach to growth curves. *Biometrika*, 62, 89-100.

62. Fuller W A (1987): Measurement Errors Models. John Wiley and Sons, New York.

63. Fisher R A (1918): On the mathematical foundation of theoretical statistics. *Trans. R. Soc.*, Edinburgh, 52, 399-433.

64. ——— (1925): Statistical Methods for Research Workers. *Oliver and Byod*, London.

65. Ganguly M (1941): A note on nested sampling. *Sankhya*, 5, 449-452.

66. Gauss K F (1809): Theoria motus corporum celestrium in sectionibus conics solem ambientium. *Pearth and Besser*, Hamburg.

67. Gelfend A E and Smith A F M (1990): Sampling based approaches to calculating marginal densities. *J A S A*, 85, 398-409.

68. Good I J (1986): The likelihood-ratio approach to the generalized Behrens-Fisher problem: one-way ANOVA under heteroscedasticity. *J S C S*, 26, 139-145.

69. Grizzle J E and Allen D M (1969): Analysis of growth and dose response curve. *Biometrics*, 25, 357-382.

70. Haslem E.(1989): Vegetables Tennins revisited, in Plant Polyphenols, *Cambridge University Press*, Cambridge, 154-219.

71. Hartley H O and Rao J N K (1967): ML estimation for the mixed analysis of variance model. *Biometrika*, 34, 233-243.

72. Harville D A (1977): Maximum likelihood approach to variance component estimation and to related problems. *J A S A*, 72, 358, 320-340.

73. Harville D A (1969a): Quadratic unbiased estimation of variance components for the one-way classification. *Biometrika*, 56, 313-326.

74. Heitjan D F (1991a): Generalized Norton-Simon models of tumor growth. *Statistics in Medicine*, 10, 1975-1988.

75. Henderson C R (1953): Estimation of variance and covariance components. *Biometrics*, 9, 213-226.

76. Herbach L H (1959): Properties of Model II type analysis of variance tests. *Annals of Mathematical Statistics*, 30, 939-959.

77. Hill B M (1965): Inference about variance components in the one-way model. *J A S A*, 60, 806-825.

78. ——— (1967): Correlated errors in the random model. *J A S A*, 62, 1387-1400.

79. Hilton P J (1989): Chemical evaluation of tea liquor. *Annual Report, Tea Research Foundation*, Central Africa, 61-63.

80. Hiller G H and Satchell S E (1986): Finite sample properties of a two-stage single equation estimator in the SUR model. *Econometric Theory*, 2, 66-74.

81. Hocking P R and Kutner M H (1975): Some analytical and numerical comparisons of estimators of the mixed AOV model. *Biometrics*, 31, 19-28.

82. Howrey E P and Varion H R (1994): Estimating the distributional impact of time-of-day pricing of electricity. *Journal of Econometrics*, 26, 12-30.

83. Jackson R W B (1939): Reliability of mental tests. *British Journal of Psychology*, 29, 267-287.

84. Jenrich R J and Schluchter D A (1986): Unbalanced repeated measures models with structured covariance matrices. *Biometrics*, 42, 805-820.

85. Jenrich R J and Sampson P F (1976): Newton-Raphson and related algorithms for maximum likelihood variance components estimation. *Technometrics*, 18, 11-17.

86. Judge GG, Hill R C, Griffiths W E and Lee T C (1985): Introduction to Theory and Practice of Econometrics. *John Wiley and Sons*, New York.

87. Kackar R N and Harville D A (1984): Approximations for standard errors of estimators of fixed and random effects in mixed linear models. *J A S A*, 79, 853-861.

88. Kenward M G (1985): The use of fitted higher-order polynomial coefficients as covariates in the analysis of growth curve. *Biometrics*, 41, 19-28.

89. Keselman H J and Lix L M (1997): Analysis of multivariate repeated measurements designs when covariance matrices are heterogeneous. *British Journal of Math. Stat. Psychology*, 50, 319-338.

90. Khatri C G (1973): Testing some covariance structures under a growth curve model. *Journal of Multivariate Analysis*, 3, 102-116.

91. Khuri A I and Sahai H (1985): Variance components analysis : A selective literature survey. *International Statistical Review*, 53, 279-300.

92. Khuri A I and Sahai H (1985): Variance components analysis : A selective literature survey. *International Statistical Review*, 53, 279-300.

93. Kleffe J and Rao J N K (1986): The extension of asymptotically unbiased non-negative quadratic estimates of variance components in ANOVA models. *J A S A*, 81, 692-698.

94. Koch G G and Landis J R (1977): A general methodology for analysis of experiments with repeated measurements of categorical data. *Biometrics*, 33, 133-158.

95. Koschat M A and Weerahandi S (1992): Chow type tests under heteroscedasticity. *J. of Business and Economic Statistics*, 10(2), 221-228.

96. Krutchkoff R G (1988): One-way fixed effects analysis of variance when error variances may be unequal. *Journal of Statistical and Computational Simulation*, 30, 259-271.

97. Kvalseth T O (1985): Cautionary note about R^2. *American Statistician*, 39, 279-285.

98. Laird N M (1982): Computation of variance components using EM algorithm. *J. Stat. Comp. & Simul.*, 14, 295-303.

99. Laird N M, Lange N and Stram D (1987): Maximum likelihood computations with repeated measures : Application of EM algorithm. *J A S A*, 82, 97-105.

100. Laird N M and Ware J H (1982): Random Effects models for longitudinal data. *Biometrics*, 38, 963-947.

101. LaMotte L R (1970): A class of estimators of variance components. *Technical Report 10*, Department of Statistics, University of Kentucky, Kentucky.

102. ——— (1971): Locally best quadratic estimators of variance components. *Technical Report 22*, Department of Kentucky, Kentucky.

103. ——— (1973a): On non-negative quadratic unbiased estimation of variance components. *J A S A*, 68, 728-730.

104. ——— (1976): Invariant quadratic estimators in the random one-way ANOVA model. *Biometrics*, 32, 793-804.

105. Lee J C (1991): Tests and model selection for the general growth curve model. *Biometrics*, 47, 147-159.

106. Legendre A M (1806): Nouvelles methodes pour la Determination des Orbites des Cometes: avec un Suppliment Contenant Divers Pefectionnements de ces Methodes et leur Application aux deux Cometes de 1805. *Courcier*, Paris.

107. Liang K Y and Zeger S L (1986): Longitudinal data analysis using generalized linear models. *Biometrika*, 73, 13-22.

108. Lindsey J K (1993): Models For Repeated Measurements. *Oxford University Press*, Oxford.

109. Lindstrom M J and Bates D M (1988): Newton-Raphson and EM algorithms for linear mixed effects models for repeated measures data. *J A S A*, 83, 1014-1022.

110. Lindley D V and Smith A F M (1972): Bayes estimates for the linear model (with discussion). *J. Royal Statist. Soc.*, Series B, 34, 1-42.

111. Madansky A (1959): The fitting of straight lines when both variables are subject to error. *J A S A*, 54, 173-205.

112. Martens M (1985): Sensory and chemical quality criteria for white cabbage studied by multivariate data analysis. *Lebensm. Wiss U. - Technol.*, 18, 100-104.

113. McDowell I, Feakes J and Gay C (1991): Phenolic composition of black tea liquors as a means of predicting country of origin. *J Sci Food Agric.*, 55, 627-641.

114. Mostler F and Tukey J W (1997): Dat Analysis and Regression : A Second Course in Statistics. *Addison-Wisley*, Reading, Massachusetts.

115. Mutsuyama Y and Ohashi Y (1997): Mixed model for bivariate response repeated measures data using Gibbs sampling. *Statistics in Medicine*, 16, 1587-1601.

116. Naes T and Solheim R (1990): Detection and interpretation of variation within and between assessors in sensory profiling. *Journal of Sensory Studies*, 6, 159-177.

117. Naes T (1998): Detecting individual differences among assessors and differences among replicates in sensory profiling. *Food Quality And Preference*, 9(3), 107-110.

118. Neyman J, Iwaszkiewicz K and Kolodziejczyk S T (1935): Statistical problem in agricultural experimentations. *J. Royal Statist. Soc.*, 2, 107-154.

119. Nummi T (1997): Estimation in a random effects growth curve model. *Journal of Applied Statistics*, 24(2), 157-169.

120. Othieno C O and Owuor P O (1984): Black tea quality and international standards. *Int. Tea J.*, 7, 27-30 .

121. Owuor P O (1995): Factory processing of black tea and its influence on tea quality in Kenya. *Tea*, 16(1), 62-69.

122. ——— (1986): Correlation of TF and TR in the evaluation of black tea. *J. Sci. Food Agric.*, 37, 507-513.

123. Owuor P O and McDowell I (1994): Changes in TF composition and astringency during black tea fermentation. *J. Sci. Food Agric.*, 37, 507-513.

124. Patthoff R F and Roy S N (1964): A generalized multivariate analysis of variance model useful for growth curve problem. *Biometrika*, 51, 313-326.

125. Pal M (1981): Estimation in Error-In-Variable Models. Ph. D. Thesis, Indian Statistical Institute.

126. Pal M, Paul S K and Das A K (1997): A Statistical approach to study the subjectivity of Taster's scores. *Two And A Bud*, 44(2), 18-25.

127. ——— (1999): Assessment of tea quality associating biochemical quality parameters with Tasters' scores". *Two And A Bud*, 45(1),

128. ———— (2000): A statistical study on the improvement of CTC tea quality through process modification. Under final review stage in *J. Sci. Food Agric.*.

129. Pal M and Paul S K (1999): Two-way heteroscedastic error component regression models : Application to tea quality assessment data. Under review in *Journal of Applied Econometrics.*

130. ———— (2000): Analysis of tea quality assessment data using heteroscedastic error component regression models. Under review in *Journal of Sensory Studies.*

131. ———— (1999): Small sample estimation problem with Box-Cox transformation model : Application to tea quality data. Under review in *Statistica Applicata.*

132. Paul S K (2000): A statistical appraisal of the problems in Tea Tasters' sensory analysis. To appear in *The Tea Research Journal.*

133. (1998): Repeated response regression model with heteroscedastic error: A study of tea quality. Paper presented in the *National Seminar on Frontiers in Statistics and Operation Research*, held in Dibrugarh University, India.

134. ———— (1999): Recent developments in the longitudinal data analysis : Linear mixed models and estimation. *Technical Report 01/99*, Department of Mathematical Sciences, Tezpur University, India.

135. ———— (2000): Detecting assessors' reproducibility under heterogeneity. Under review in *Food Quality and Preference.*

136. Patterson H D and Thomson R (1971): Recovary of inter-block information when block sizes are unequal. *Biometrika*, 58, 545-554.

137. ———— (1974): Maximum likelihood estimation of components of variance. *Proceedings of Eighth International Biometric Conference*, 197-209.

138. Puri M L and Sen P K (1985): Nonparametric Methods in Multivariate Analysis. *John Wiley*, New York.

139. Prentice R L and Zhao L P (1991): Estimating equations for parameters in means and covariances of multivariate discrete and continuous responses. *Biometrics*, 47, 825-839.

140. Rao C R (1973): Linear Statistical Inference and its Applications, Wiley Eastern Ltd., New Delhi.

141. ——— (1965): The theory of LS when the parameters are stochastic and its application to the analysis of growth curves. *Biometrika*, 52, 447-458.

142. ——— (1972): Estimation of variance and covariance components : MINQUE theory. *J. Multivariate Analysis*, 1, 257-275.

143. ——— (1970): Estimation of heteroscedastic variances in linear models. *J A S A*, 65, 161-172.

144. Rao P S R S, Kaplan J and Cochran W G (1981): Estimation of one-way random effects model with unequal error variances. *J A S A*, 76, 89-96.

145. Roberts G R and Fernado R S S(1981): Some observations on the correlation of polyphenol content to quality of tea clones", *Tea Qurt.*, 50, pp.30-32.

146. Rosen D (1991): The growth curve model : A review. *Comm. in Statistics, Theory and Methods*, 20, 2791-2822.

147. Roy S N (1967): The heuristic method of test construction and its use in multivariate analysis. *Annals of Math. Stat.*, 24, 220-238.

148. Russel T S and Bredley R A (1958): One way variances in the two-way classification. *Biometrika*, 45, 111-129.

149. Ryan L M and Dempster A P (1984): Weighted Normal Plots. *Technical Report 394Z*, Farber cancer Institute, Boston.

150. Scheffe H (1956): Alternative models for the analysis of variance. *Ann. Math. Stat.*, 27, 251-271.

151. ———— (1959): The Analysis of Variance. *John Wiley and Sons*, New York.

152. Searle S R (1971): Linear Models. *John Wiley and Sons*, New York.

153. Searle S R and Henderson H V (1979): Dispersion matrices for variance components models. *J A S A*, 74, 465-470.

154. Searle S R , Casella G and McCulloch R J (1992): Variance Components. *John Wiley and Sons*, New York.

155. Smith A F M (1983): Discussion on an article by DuMouchel et al., *J A S A*, 78, 310-311.

156. Solomon P J and Cox D R (1992): Nonlinear component of variance models. *Biometrika*, 79, 1-11.

157. Stanek E.J. III and O'Hearn J.R.(1998) : "Estimating Realized Random Effects", *Commun. Statist.- Theory Meth.*, 27(5), 1021-1048.

158. Tan W Y and Tabatabai (1986): Some Monte Carlo studies on the comparison of several means under heteroscedasticity and robustness with respect to depurtures from normality. *Biometrical Journal*, 28, 801-810.

159. Temple C M (1996): Simple chemical indicators of value of black tea", *Tea Research Foundation Quarterly News Letters*, 122, pp.43-48.

160. Thomson W A (1962): The problem of negative estimates of variance components. *Ann. Math. Stat.*, 33, 273-289.

161. Tiao G C and Tan W Y (1965): Bayesian analysis of random effects models in analysis of variance I : Effects of autocorrelated errors. *Biometrika* , 52, 37-53.

162. Tiao G C and Tan W Y (1966): Bayesian analysis of random effects models in analysis of variance II : Effects of autocorrelated errors. *Biometrika* , 53, 477-499.

163. Tierney L and Kadane J B (1986): Accurate estimation for posterior moments and marginal densities. *J A S A*, 81, 82-86.

164. Tippett L H C (1931): The Methods of Statistics. *Willums and Norgate*, London.

165. Townsend E C (1968): Unbiased estimators of variance components in simple unbalanced designs. Ph. D. Thesis, Carnell University, New York.

166. Townsend E C and Searle S R (1971): BQUE of variance components from unbalanced data in one-way classification. *Biometrics*, 27, 643-657.

167. Tunnichiffe-Wilson G.(1989): "On the use of Marginal Likelihood in Time Series Model Estimation", *J. of Royal Stat. Soc.*, B 51, 15-27.

168. Tukey J W (1977): Exploratory Data Analysis. *Addison-Wisley*, Reading.

169. Ullah M R (1989): A rapid procedure for estimating Theaflavins and Thearubigins of black tea", *Two and a Bud*, 33, pp.46-48.

170. Verbyla A P and Cullis B R (1990): Modelling in repeated measures experiments", *Applied Statistics*, 39, 341-356.

171. Verbyla A P (1986): Conditioning in the growth curve model. *Biometrika*, 73, 475-483.

172. Verbyla A P and Vanebles W N (1986): An extension of the growth curve model. *Biometrika*, 76, 129-138.

173. Vonesh N J and Chinchilli R P (1997): Linear and Nonlinear Models for The Analysis of Repeated Measurements. *Marcel Dekker, Inc.*, New York.

174. Vonesh H J, Chinchilli R P and Pu K (1996): Goodness-of-fit in generalized nonlinear mixed effects models. *Biometrics*, 52, 572-587.

175. Wansbeek T and Kapteyn A (1982): A simple way to obtain spherical decomposition of variance components models for balanced data. *Coommun. Statist. - Theory and Methods*, 11(18), 2105-2112.

176. Ware J H (1985): Linear models for the analysis of longitudinal studies. *American Statistician*, 39, 95-101.

177. Wald A (1940): A note on the analysis of variance with unequal class frequencies. *Annals of Math. Stat.*, 11, 96-100.

178. Welch B L (1951) On the comparison of several means values : An alternative approach. *Biometrika*, 38, 330-336.

179. Weerahandi S (1991): Testing variance components in mixed models with generalized p values. *J A S A*, 86, 151-153.

180. Winsor C P and Clarke G L (1940): Statistical study of variation in the catch of plankton nets. *Sears Foundation J. marine*, 3, 1-34.

181. Yanez N.D. III, Kronmal R.A. and Shemanski L.R. (1998): "The Effects of Measurement Error in Response Variables and Teasts of Association of Explanatory Variables in Change Models", *Statistics in Medicine*, 17, 2597-2606.

182. Yamanishi J (1995): Flavour of tea. *Food Review International*, 11(3), 477-525.

183. Zellner A (1962): A efficient method of estimating seemingly unrelated regression and tests for aggregation bias. *J A S A*, 57, 348-368.

184. Zellner A and Huang D S (1962): Further properties of efficient estimators for seemingly unrelated regression equations. *International Economic Review*, 3, 300-313.

185. Zhou L and Methew T (1994): Some tests of variance components using generalized p values. *Technometrics*, 36, 394-402.