

RESTRICTED INFORMATION

**SOME RESULTS CONCERNING ASYMPTOTIC DISTRIBUTIONS
AND THEIR APPLICATIONS**

RESTRICTED
SECTION

**SOME RESULTS CONCERNING ASYMPTOTIC DISTRIBUTIONS
AND THEIR APPLICATIONS**

By

J. SETHURAMAN

RESEARCH AND TRAINING SCHOOL

INDIAN STATISTICAL INSTITUTE

CALCUTTA

1961

P R E F A C E

This thesis is being submitted to the Indian Statistical Institute in support of the author's application for the degree of Doctor of Philosophy. The thesis embodies research carried out by the author during the period 1958-1961 under the supervision of Dr. R. R. Bahadur, Professor of Statistics at the Indian Statistical Institute, Calcutta.

This thesis is concerned with the development of a new method of establishing asymptotic distributions and the elaboration of some of its applications in several fields of Statistics.

The thesis consists of five Chapters. Chapter I describes, in general terms, the various problems considered in the thesis. Chapter II deals with the basic results concerning the convergence of joint distributions under assumptions about the convergence of marginal and conditional distributions. Chapter III is devoted to the applications of the results of Chapter II to some problems in Statistical Methods. Chapters IV and V deal with some applications to Regression Analysis. Chapter V also contains the results of some model sampling experiments, and tables for practical use.

Earlier versions of some of the results of Chapter III have appeared as a joint article with Dr. B. V. Sukhatme of the Indian Council of Agricultural Research, New Delhi, in Sankhya, 21 (1959) 289-298 (Ref. [49]). Some of the results of Chapter IV have appeared in Sankhya, Series, A, 23(1961), 73-90. (Ref. [48]).

The author wishes to express his sincere gratitude to Dr. R.R. Bahadur for his guidance and advice on the method of presentation of the material in the thesis, and to Dr. C. R. Rao, Professor and Head of the Division of Theoretical Research and Training, Indian Statistical Institute, for his constant encouragement.

The author is indebted to the following members of the Indian Statistical Institute: to Sri Nikilesh Bhattacharya for his help in some model sampling experiments and for his reading Chapters IV and V; to Sri G. Parthasarathy for his help in some model sampling experiments; and to Sri T. Krishnamurthi and to Sri Debdas Choudhuri for their computation of tables I, II, III and IV on the electronic computer. The author also wishes to express his gratitude to Sri K.R. Parthasarathy of the Indian Statistical Institute for procuring and translating N.V. Smirnov's article (Ref. [50]), which necessitated a revision of an earlier version of Chapter III.

The author also records his gratefulness to the Research and Training School of ^{The} Indian Statistical Institute for providing facilities for research. Thanks are due to Sri G.M. Das for his efficient typing of the Thesis.

C O N T E N T S

	<u>Page</u>
CHAPTER I	
INTRODUCTION	1 - 11
1.1. The technique of conditional distributions.	1
1.2. Applications to Statistical Methods.	5
1.3. Applications to Regression Analysis.	6
1.4. Brief summary of the Chapters of the Thesis	9
CHAPTER II	
MAIN THEOREMS	12 - 28
2.1. Summary and Introduction.	12
2.2. Notations and Preliminaries.	12
2.3. Main Theorems.	17
2.4. A counterexample.	26
2.5. The use of theorems 3.2, 3.3, 3.4 and 3.9.	26
CHAPTER III	
APPLICATIONS TO STATISTICAL METHODS	29 - 70
3.1. Introduction and summary.	29
3.2. Notations, definitions and basic preliminaries.	31
3.3. Auxiliary Theorems.	43
3.4. Asymptotic distributions of U-statistics and order statistics.	47
3.5. Corollaries of Theorems in section 4.	56
3.6. Asymptotic distributions of V-statistics and order statistics.	59

CHAPTER III (Contd.)

- 3.7. Alternative proofs and generalisations of the theorems of Hoeffding concerning mixtures of order statistics. 64

CHAPTER IV

FIXED INTERVAL ANALYSIS AND FRACTILE ANALYSIS - PART I 71 - 118

- 4.1. Introduction and summary. 71
- 4.2. The methods of fixed interval analysis and fractile analysis. 73
- 4.3. Notations, definitions, etc. 85
- 4.4. Limit distributions. 91
- 4.5. Theoretical applications. 104
- 4.6. A general formulation of the regression problem. 111

CHAPTER V

FIXED INTERVAL ANALYSIS AND FRACTILE ANALYSIS - PART II. 119-157

- 5.1. Introduction and summary. 119
- 5.2. Testing in fixed interval analysis. 121
- 5.3. Testing in fractile analysis. 123
- 5.4. Model sampling for limit distributions in fixed interval analysis. 131
- 5.5. Model sampling experiments in fractile analysis. 132

CHAPTER V (Contd.)

TABLES		135-157
Table I:	(Q_g) for $g = 2(1)16, 20, 25$.	135
Table II:	$(q_{g,1}, \dots, q_{g,g})$ for $g = 2(1)16, 20, 25$.	148
Table III:	$\text{Tr } Q_g$ for $g = 2(1)16, 20, 25$.	151
Table IV:	$\sum_1^g \mu_{g,i}^2$ for $g = 2(1)16, 20, 25$.	152
Table V:	Model sampling experiments in fixed interval analysis.	153
Table VI:	Model sampling experiments in fractile analysis.	154
Table VII:	Model sampling experiments in fractile analysis.	155
Table VIII:	Model sampling experiments in fractile analysis.	156
Table IX:	Model sampling experiments in fractile analysis.	157
LIST OF REFERENCES		158-162

Chapter I.

INTRODUCTION

1.1. The technique of conditional distributions.

The theory of limit theorems for random variables occupies a very important position in probability theory. In Statistics it goes under the name asymptotic distribution theory. The parameter, n , that tends to infinity in these limit theorems is usually the sample size and in those instances where small sample distributions are difficult to obtain or compute, the asymptotic distributions which are easier to obtain serve as convenient approximations. This fact alone is enough to emphasise their usefulness in statistics.

To define convergence and limit distributions of random variables we should first examine how a distribution is defined. If there are several ways of defining a distribution, a natural mode of convergence can be associated with each definition. The study of the interrelation of these different modes of convergence will be useful in deciding whether a given sequence of distributions converges in an 'accepted' mode.

Let us, for the purposes of elaborating our point, assume that our sample space \mathcal{Z} is $\mathcal{X} \times \mathcal{Y}$ where \mathcal{X} and \mathcal{Y} are the real line.

The field \mathcal{L} , of measurable sub sets of Z is $\mathcal{A} \times \mathcal{B}$ where \mathcal{A} and \mathcal{B} are the usual fields of Borel subsets on the real line. A random variable ζ taking values in Z will be of the form (ξ, η) , where ξ takes values in X and η in Y . By the distribution of ζ we mean a probability measure $\lambda(\cdot)$ defined over the field \mathcal{L} . The 'accepted' mode of convergence of a sequence ζ_n of random variables to ζ , known as weak convergence, is defined as follows:

ζ_n converges weakly to ζ if and only if

$$\int_Z g d\lambda_n \rightarrow \int_Z g d\lambda \quad \dots \quad (1)$$

for every bounded continuous function g on Z .

A distribution can be defined through a function $F(x, y)$, called the distribution function, with the following properties.

$$0 \leq F(x, y) \leq 1$$

$$F(x + h, y + k) - F(x + h, y) - F(x, y + k) + F(x, y) \geq 0$$

$$\text{for all } h, k \geq 0$$

$$F(-\infty, y) = F(x, -\infty) = 0 \quad \text{for all } -\infty < x, y < \infty.$$

$$F(\infty, \infty) = 1.$$

(2)

The process of constructing a probability measure λ from such a function $F(z)$ is well known. See for instance Cramer (1946). $\{\xi_n\}$ could now be said to converge to ζ if and only if

$$F_n(z) \rightarrow F(z) \quad \text{for each } z \quad \dots \quad (3)$$

that is a continuity point of $F(z)$. It is well known (see for instance Cramer (1946)) that definitions (1) and (3) are equivalent.

A distribution can also be studied in terms a function called its characteristic function. It is known that a characteristic function determines a distribution uniquely. Thus we may now ask ourselves: 'If a sequence of characteristic functions $\varphi_n(t)$ converges to $\varphi(t)$ pointwise, will the corresponding sequence of random variables converge weakly?' The famous continuity theorem (see Levy (1937), Cramer (1937)) answers this question in the affirmative under the condition that $\varphi(t)$ is continuous at 0.

Another method of defining a distribution is to provide all the moments ($\alpha_{10}, \alpha_{01}, \alpha_{20}, \dots$) of the distribution. (We can discuss only distributions with all moments by this method.) Whether this method is successful or not is the famous moment problem and a sufficient condition for success has been given by Carleman. See for instance Kendall and Stuart (1958). We have the following theorem (due to

4

Fréchet and Shohat) connecting a definition of convergence through moments with weak convergence; (see Kendall and Stuart (1958)).

If the moments sequence $(\alpha_{10}^{(n)}, \alpha_{01}^{(n)}, \alpha_{20}^{(n)}, \dots)$ of ζ_n converges to a sequence $(\alpha_{10}, \alpha_{01}, \alpha_{20}, \dots)$ term by term and if $(\alpha_{10}, \alpha_{01}, \alpha_{20}, \dots)$ determines the distribution function of a random variable ζ uniquely then ζ_n converges to ζ weakly.

It is also known that a vector random variable ζ is uniquely determined by the collection of the unidimensional distributions of all the linear functionals $L\zeta$ of ζ . See for instance Cramer (1946). Varadarajan (1958) has shown that if for every linear functional L , $L\zeta_n$ converges weakly to a random variable ζ_L then ζ_n converges weakly to a random variable ζ and $\zeta_L = L\zeta$.

Yet another method of defining the distribution of ζ is to define the marginal distribution of ζ and, for each x , the conditional distribution of η given $\zeta = x$. The question that is raised in this work is the following: 'If the marginal distributions of ζ_n converge (in some sense) to a distribution, and the conditional distributions of η_n given ζ_n converge, (again, in some sense) can we assert that ζ_n will converge weakly to a random variable ζ and further that the distribution of ζ can be constructed in the natural way from the limiting marginal and conditional distributions?' In chapter I we give precise formulations of this question, and provide some answers. We also show by means of a counter example that unless the convergences are suitably

restricted, the answer is not in the affirmative in general.

A sort of converse to our problem has been discussed in the literature. The converse is, roughly speaking, "if ξ_n converges to ξ weakly, can we assert that η_n given $\xi_n = x$ converge weakly to η given $\xi = x$ for each x ?" This question arose from the need to construct a rigorous version of Fisher's proof of the distribution of the discrepancy χ^2 . See Fisher (1922). Several authors have discussed this question among whom we may mention Steak (1957) and Gihman (1953).

1.2. Applications to Statistical Methods.

One of the common problems in asymptotic distribution theory is the evaluation of the asymptotic distribution of a statistic based on a sample of n independent and identically distributed observations, where n is allowed to tend to infinity. The origins of the central theorem can be traced to one such problem where the statistic concerned was the sample mean. ~~Heoffding in~~ Heoffding (1948) defined a statistic called a U-statistic which is a generalisation of the sample mean, and showed that it has an asymptotic normal distribution (v. Mises ⁽¹⁹⁴⁷⁾ has also shown this in ~~V. Mises (1947)~~). Another important class of statistics whose asymptotic distributions have been investigated is the class of order statistics. Fréchet (1927), Fisher and Tippot (1928), Gumbel (1958) Gnedenko (1943)(whose results are complete and exhaustive) and others have found the asymptotic distribution of the extreme value, i.e. the largest or the smallest order statistic. A proof that under certain

certain conditions the sample quantile is asymptotically normally distributed is found in Granger (1946). By far the most complete and exhaustive treatment of the asymptotic distributions of order statistics is due to Smirnov (1952). He has enumerated all possible types of limiting distributions for some broad groups of order statistics. Dworetz (1951) has evaluated the asymptotic distribution for another group of order statistics.

We observe that U-statistics and order statistics have certain similarities; for instance they are symmetric functions of the observations; but their asymptotic behaviours are widely different. It is thus of interest to investigate the nature of the asymptotic joint distributions of U-statistics and order statistics. This investigation is carried out in Chapter III with the help of the techniques of Chapter II. Earlier work on this problem are found in Sukhatme (1957), and Sukhatme and the author (1959). In these papers, the use of the technique of characteristic functions has led to the imposition of many restrictions, which are removed in the present treatment. Another class of statistics called V-statistics, which are generalisations of U-statistics, is introduced in Chapter III and the asymptotic joint distributions of V-statistics and order statistics are investigated.

1.3. Applications to Regression Analysis.

Chapters IV and V contain contributions to regression analysis. One of the important problems of Statistics is the study of the pattern of relationship of one variable Y with another variable

X and a comparison of such relationships. These are usually done through a study of the regression function $\lambda(x)$ of Y on X ($\lambda(x) = E(Y | X = x)$.) For the comparison of $\lambda(x)$ and $\lambda'(x)$, the regression functions in two populations P and P', the methods adopted in practice at present, are those that assume that $\lambda(x)$ and $\lambda'(x)$ are of some convenient algebraic form completely determined except for a finite number of parameters. Thus the classical method of linear estimation and least squares assumed $\lambda(x)$ to be a linear function of x. Other methods have assumed $\lambda(x)$ to be a polynomial in x, or a polynomial in sines and cosines of x, or of the logistic distribution form, or of the normal probability function form, or an exponential polynomial, etc. An advantage of these methods is that they reduce the whole problem to one of estimating and testing a finite number of parameters. Elementary accounts of these methods are readily available; for instance see Williams (1959), Plackett (1960). A disadvantage of these methods is that in many instances one of these regression models will have to be assumed without reference to the practical conditions obtaining in the problem. Further, there are no satisfactory techniques to find out which model we are to choose in a given situation.

We show in Chapters IV and V how we can, in a very simple way, do without these cumbersome regression models in many practical situations. The method consists in comparing, by a non-parametric method, the regressions $\bar{y}(a, b)$, $\bar{y}'(a, b)$ over intervals

($\bar{Y}(a, b) = E(Y|a < X \leq b)$) rather than the pointwise regressions $\lambda(x)$, $\lambda'(x)$. In many instances the intervals (a, b) can be chosen carefully to make the comparisons meaningful. This method is known as the method of fixed interval analysis.

There are problems where it is difficult to demarcate intervals for the two variables X and X' that would be comparable in the sense we have in mind. For instance, if X and X' represent the total expenditures of individuals in two countries with different currencies, it is difficult to assign intervals for X and X' that would correspond to similar socio-economic classes in the two countries. But it can be safely assumed that the total expenditure is a monotonic function of the socio-economic level of a population. Thus if the values of X and X' are grouped into classes on the basis of their ranks, these classes should be comparable. A comparison of the regressions in these classes is thus meaningful and can be made. Professor P. C. Mahalanobis was the first to note that there are instances such as the above where fixed interval analysis cannot be used. The modification suggested above was given by him and is known as fractile analysis. See Mahalanobis (1958a), (1958b), (1960).

In practice some suitable method, for instance the method explained in Chapters IV, V can be adopted to obtain estimates of the regression function, in fixed interval analysis and in fractile analysis. The equality of the two regressions is then decided by the use of some

suitable measure of divergence based on these estimates. To perform the test the sampling distributions of these statistics are needed. In Chapter IV we evaluate the asymptotic distribution of these statistics (with the use of the theorems of Chapter II) and show that in large samples the problem of testing the equality of regressions is equivalent to testing the equality of means in two multivariate normal populations. According to this reduction each sample yields only one observation from a multivariate normal distribution and so the reduction is useful only when several independent and equally valid samples are available from each population.

In Chapter V we discuss other methods that can be used when just one sample is available from each population. While using the method of fractile analysis if we assume that (Y, X) is bivariate normal then we can make use of the tables I, II, III and IV in Chapter V for performing tests for the equality of regressions in large samples.

1.4. Brief summary of the Chapters of the thesis.

Chapter II. Main Theorems.

This chapter contains the main theorems of the thesis. Theorems 2.3.2, 2.3.3, 2.3.4, 2.3.9 show that under certain conditions the convergence of the marginal and conditional distributions implies the convergence of the joint distributions. Section 2.4 provides a counter example which demonstrates that some restrictions like the ones in theorems

2.3.2, 2.3.3, 2.3.4, 2.3.9 are in general necessary.

Chapter III. Applications to Statistical Methods.

This chapter contains applications of the theorems of Chapter II in obtaining certain asymptotic distributions. Theorems 3.4.1, 3.4.3, 3.4.5 establish the only possible limit laws for a U-statistic and an order statistic belonging to one of three major groups. Theorems 3.6.1, 3.6.2 concern similar results ^{for} V-statistics and order statistics. 3.7 contains some simple proofs and generalisations of Hoeffding's theorems on mixtures of order statistics. These results are useful in a discussion in Chapter V.

Chapter IV. Fixed Interval Analysis and Fractile Analysis-Part I.

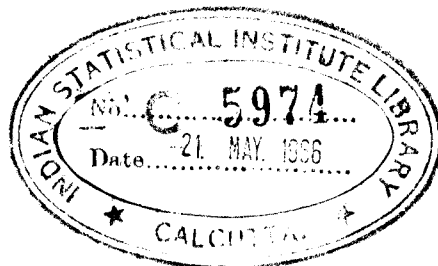
Section 4.2 contains a detailed exposition of the methods of fixed interval analysis and fractile analysis. Limit Theorems required in this connection are studied in Section 4.4. Section 4.5 explains how in large samples the general problem can be reduced to one of testing the equality of the means of two multivariate normal populations. In section 4.6 the methods of fixed interval analysis and fractile analysis are formulated in a general way and some new unsolved problems are posed.

Chapter V. Fixed Interval Analysis and Fractile Analysis - Part II.

The chapter begins with a description of the practical techniques that can be adopted when only one sample is available from each population. Tables are provided for estimating certain quantities that are

involved in the test function of fractile analysis when (Y, X) is assumed to be bivariate normal.

The results of some model sampling experiments are given, which give an empirical idea of the rate at which the limiting distributions are reached as the sample size increases.



Chapter II

MAIN THEOREMS

2.1. Summary and Introduction.

In this chapter we present our fundamental theorems. We have reviewed, in Chapter I, several techniques of finding the asymptotic distributions of statistics. We develop in section 3 a new method for obtaining the joint asymptotic distribution of several statistics when some information relating to the marginal and conditional distributions is available.

When Z is the product space $X \times Y$, a random variable ζ on (Z, U) is of the form (ξ, η) where ξ is a random variable on (X, S) and η on (Y, T) . Let $\{\zeta_n\}$ be a sequence of random variables on Z and let $\lambda_n(\cdot)$, $\mu_n(\cdot)$ and $\nu_n(x, \cdot)$ denote the joint distribution of (ξ_n, η_n) , the marginal distribution of ξ_n and the conditional distribution of η_n given $\xi_n = x$, respectively. The theorems of this chapter are concerned with the convergence of $\{\lambda_n\}$ when $\{\mu_n\}$ and $\{\nu_n\}$ are known to converge in some sense.

2.2. Notations and Preliminaries.

Before embarking on the statement and the proofs of our theorems, we explain in this section our notations and mention some well known results which form the basic tools of this chapter.

Throughout this chapter we will be concerned with two measure spaces (X, S) and (Y, T) , their product $(Z, U) = (X \times Y, S \times T)$ and a sequence of random variables (ξ_n, η_n) taking values in Z . The distribution of (ξ_n, η_n) on (Z, U) will be denoted by λ_n while the distribution of ξ_n on (X, S) will be denoted by μ_n . We assume that the conditional distribution of η_n given $\xi_n = x$ exists as a probability measure, i.e. there exists a function $\nu_n(x, B)$ which is a probability measure on T for each $x \in X$ and is a measurable function on X for each $B \in T$ and further satisfies the equation

$$\lambda_n(A \times B) = \int_A \nu_n(x, B) d\mu_n \quad \text{for all } A \in S \text{ and } B \in T$$

In this case, if C is any set in U and C_x for each $x \in X$ denotes the subset $\{y : y \in Y, (x, y) \in C\}$, then

$$\lambda_n(C) = \int_X \nu_n(x, C_x) d\mu_n \quad (\text{See Halmos (1950)})$$

We shall also denote $\lambda_n(A \times B) / \mu_n(A)$ by $\nu_n(A, B)$ with an obvious interpretation.

If p_1, p_2, \dots is a sequence of probability measures defined on a measurable space (M, V) , we shall say p_n converges strongly to p ($p_n \rightarrow p$ in symbols) if $p_n(C) \rightarrow p(C)$ for each $C \in V$. It is well known that for a given sequence $\{p_n\}$ there exists

a p such that $p_n \rightarrow p$ if and only if $\lim_n p_n(C)$ exists for all $C \in \mathcal{V}$; in this case the p_n 's are equicontinuous, i.e. if $\{C_k\}$ is any decreasing sequence of sets with $\bigcap_{k=1}^{\infty} C_k = \emptyset$, then $\sup_n p_n(C_k) \rightarrow 0$ as $k \rightarrow \infty$. (See Halmos (1950)). It is also known that $p_n \rightarrow p$ if and only if $\int g dp_n \rightarrow \int g dp$ for all bounded measurable functions g . (See Halmos (1950)). If the densities $f_n(m)$ of p_n with respect to some σ -finite measure p_0 converges in measure $[p_0]$ to a density function $f(m)$ then there is a p such that $p_n \rightarrow p$. (See Scheffe (1947)). The above is a sufficient condition for the strong convergence of a sequence of probability measures $\{p_n\}$ which is convenient in practice.

We shall also require the notion of weak convergence of probability measures. This requires that the basic space be topological, and that all continuous functions be measurable. If p_1, p_2, \dots , are probability measures on such a space M , we shall say p_n converges to p weakly ($p_n \Rightarrow p$ in symbols) if $\int g dp_n \rightarrow \int g dp$ for all bounded continuous functions $g(m)$ on M . A set C is said to be a continuity set of p if $p(\text{bd } C) = 0$ where $\text{bd } C$ is the boundary of C . $p_n \Rightarrow p$ if and only if $\lim_n p_n(C) = p(C)$ for each C that is a continuity set of p . See Billingsley (1956). Further, if M is separable complete metric, $\{p_n\}$ is compact under weak convergence if and only if, for each $\epsilon > 0$, there is a compact set $C \subset M$ with $p_n(C) \geq 1 - \epsilon$ for all n . See Prohorov (1956), Varadarajan (1958a).

We will also have occasion to deal with a metric space X with a metric d . The σ -field S on X will be taken to be the Borel field. We denote the diameter of any subset F of X by $\delta(F)$. The subset $S(F, \delta)$ denotes the set $\{x: d(x, x_1) < \delta, x_1 \in F\}$.

We say that a family \mathcal{F} of sets is a Vitali covering of a subset K of X with respect to a measure μ on S if

- i) $F \in \mathcal{F}$ implies that F is closed and $\mu(F) > 0$.
- ii) Given any $\epsilon > 0$ there is a fixed number α such that $x \in K$ implies that there is an $F \in \mathcal{F}$ such that $x \in F$, $\delta(F) < \epsilon$ and $\mu(S(F, 3\delta(F))) / \mu(F) \leq \alpha$. Actually, here we can take $2 + \beta$ with $\beta > 0$ instead of 3. For theorems and definitions related to Vitali covering refer to Dunford and Schwartz (1958), Munroe (1953) etc. We recall here a well known theorem due to Vitali.

2.2.1. Theorem (Vitali)

If a family of sets \mathcal{F} is a Vitali covering of a compact set K of a metric space X with respect to a measure μ then there exists a sequence of sets $\{F_n\} \subset \mathcal{F}$ such that

- i) the sets F_n are disjoint
- ii) $\mu\left[K - \bigcup_{n=1}^{\infty} F_n\right] = 0$.

The theorem can be equivalently expressed as follows: Given any $\epsilon > 0$, there is a finite collection of sets F_1, \dots, F_m in such that $\mu\left[K - \bigcup_{n=1}^m F_n\right] < \epsilon$.

For the formulation of one of our theorems we require the notion of UC* convergence (allied to that of Parzen (1954)) of a family of sequences of probability measures. Let $\gamma_n(\theta, \cdot)$, $n = 0, 1, \dots$ be a family of sequences of probability measures on $M = R_k$, the euclidean space of k dimensions. It is assumed that the index θ takes values in a compact metric space I . Let $\varphi_n(t, \theta)$ denote the characteristic function of $\gamma_n(\theta, \cdot)$, i.e.

$$\varphi_n(t, \theta) = \int_M \exp(i \underline{t} \underline{m}') \gamma_n(\theta, d\underline{m}), \quad n = 0, 1, \dots \quad (1)$$

$\gamma_n(\theta, \cdot)$ is said to converge to $\gamma_0(\theta, \cdot)$ in the UC* sense relative to $\theta \in I$ if

$$\text{a) } \sup_{\theta \in I} |\varphi_n(t, \theta) - \varphi_0(t, \theta)| \rightarrow 0 \text{ as } n \rightarrow \infty \quad \left| \dots \right. \quad (2)$$

and

$$\varphi_0(t, \theta) \text{ is equicontinuous in } \theta \text{ at } t = 0 \quad \left| \dots \right.$$

and

$$\text{b) } \varphi_0(t, \theta) \text{ is a continuous function of } \theta \text{ for each } t. \quad \dots (3)$$

(2) alone forms the definition of Parzen (1954). (2) and (3)

imply that $\varphi_0(t, \theta)$ is continuous in t and θ . (2) and (3) also imply the following

$$\theta_n \rightarrow \theta_0 \text{ implies } \varphi_n(t, \theta_n) \rightarrow \varphi_0(t, \theta_0) \dots \quad (4)$$

which means

$$\gamma_n(\theta_n, \cdot) \rightarrow \gamma_0(\theta_0, \cdot) \dots \quad (5)$$

$$\theta_n \rightarrow \theta_0 \text{ implies } \varphi_0(t, \theta_n) \rightarrow \varphi_0(t, \theta_0) \dots \quad (6)$$

which means

$$\gamma_0(\theta_n, \cdot) \rightarrow \gamma_0(\theta_0, \cdot) \dots \quad (7)$$

2.3. Main Theorems.

Before proceeding to state and prove the first of our main results we establish the following lemma.

2.3.1. Lemma.

Let $\{p_n\}$ be a sequence of probability measures on an arbitrary measure space (M, \mathcal{V}) converging strongly to a measure p . Let $\{u_n\}$ be a sequence of uniformly bounded \mathcal{V} -measurable functions converging almost everywhere $[p]$ to a function u . Then

$$\int_M u_n dp_n \rightarrow \int_M u dp.$$

Proof:

Let $|u_n(m)| \leq A$ for all n . Define

$$D_n = M - \bigcup_{k=1}^n \left\{ m : |u_k(m) - u(m)| < \epsilon \text{ for } k = n, n+1, \dots \right\}$$

We know that the D_n are decreasing and that if $\bigcap D_n = D$ then $\mu(D) = 0$.

We have

$$\begin{aligned}
& \left| \int_M u_n dp_n - \int_M u dp \right| = \left| \int_M (u_n - u) dp_n + \int_M u dp_n - \int_M u dp \right| \\
& \leq \left| \int_{D_n} (u_n - u) dp_n \right| + \left| \int_{M-D_n} (u_n - u) dp_n \right| + \left| \int_M u dp_n - \int_M u dp \right| \\
& \leq 2A p_n(D_n) + \epsilon + \left| \int_M u dp_n - \int_M u dp \right| \dots \tag{8}
\end{aligned}$$

The first term tends to zero since p_n are equicontinuous and the last term tends to zero since $p_n \rightarrow p$. Since $\varepsilon > 0$ is arbitrary, the Lemma is proved.

Suppose that, in some sense, the (marginal) distributions μ_n converge to μ and the (conditional) distributions $\nu_n(x, \cdot)$ converge to $\nu(x, \cdot)$. Further, let the joint distributions λ_n converge. It is plausible that λ_n then converges to the distribution λ_0 defined by μ and $\nu(\cdot, \cdot)$, i.e.

$$\lambda_0(A \times B) = \int_A \nu(x, B) d\mu \dots \quad (9)$$

over rectangle sets $A \times B$. This defines a distribution λ_0 uniquely on (Z, U) . In what follows, by λ_0 we mean the distribution defined by the relation (9).

2.3.2. Theorem.

If the sequence of (marginal) distributions $\{\mu_n\}$ converges strongly to μ and if for almost all x $[\mu]$ the sequence of (conditional) distributions $\{\nu_n(x, \cdot)\}$ converges strongly to $\nu(x, \cdot)$ then the sequence of (joint) distributions $\{\lambda_n\}$ converges strongly to λ_0 .

Proof:-

Let $g(x, y)$ be any U -measurable function on Z , bounded by K . We define sequence of S -measurable functions $\psi_n(x)$ as follows:

$$v_n(x) = \int_Y g(x, y) \nu_n(x, dy) \dots \quad (10)$$

It is plain that $|v_n(x)| \leq K$ and that $v_n(x) \rightarrow v(x) =$

$$v_n(x) \rightarrow v(x) = \int_Y g(x, y) \nu(x, dy) \dots \quad (11)$$

for almost all $x [\mu]$.

On application of lemma 3.1 we find that

$$\begin{aligned} \int \int g(x, y) d\lambda_n &= \int_X d\mu_n \int_Y g(x, y) \nu_n(x, dy) \\ &= \int_X v_n(x) d\mu_n \rightarrow \int_X v(x) d\mu \\ &= \int_X d\mu \int_Y g(x, y) \nu(x, dy) \\ &= \int \int g(x, y) d\lambda \dots \end{aligned} \quad (12)$$

where

$$\lambda(C) = \int_X \nu(x, C_x) d\mu, \dots \quad (13)$$

$C_x = \{y: y \in Y, (x, y) \in C\}$ - (See Halmos (1950)).

This λ is the same as λ_0 . Thus $\lambda_n \rightarrow \lambda_0$.

2.3.3. Theorem.

If the sequence of (marginal) distributions $\{\mu_n\}$ converges strongly to μ and if the sequence of (conditional) distributions $\{\nu_n(x, \cdot)\}$ converges weakly to $\nu(x, \cdot)$ for almost all $x[\mu]$, then the sequence of (joint) distributions $\{\lambda_n\}$ converges weakly to λ_0 .

Proof:

The proof of this theorem is on the same lines as theorem 3.2 and so is omitted.

In theorems 3.2, 3.3 we assumed that the marginal distributions converge strongly. We now ask ourselves what happens if the marginal distributions converge only weakly. Naturally, one expects that one shall have to strengthen the mode of convergence of the conditional distributions. We have given an example in section 4 to illustrate this fact, that, in general, such strengthening would be necessary. The difficulty in this situation is that the conditional distribution at the n th stage is defined almost everywhere with respect to μ_n and the $[\mu_n]$ null x -sets, the sets of misbehaviour of $\nu_n(x, \cdot)$, vary with n . Thus we should introduce some smoothness restriction on the conditional distributions.

For the purposes of the following theorem we will restrict X to be a locally compact separable metric space and Y to be a separable complete metric space.

2.3.4. Theorem.

Let the (marginal) distributions μ_n converge weakly to a distribution μ . Let K be the smallest closed set in X with $\mu(K) = 1$. Let \mathcal{F} be a family of continuity sets of μ covering K in the Vitali sense with respect to μ . Let the (conditional) distributions $\nu_n(F, \cdot)$ converge weakly to a distribution $\nu(F, \cdot)$ for each $F \in \mathcal{F}$. Then the (joint) distributions λ_n converge weakly

to a distribution λ .

Moreover if $\nu(A_n, B) \rightarrow \nu(x_0, B)$ where $x_0 \in A_n$, A_n is closed and $\nu(A_n) \rightarrow 0$ (for almost all $x_0 \in [\mu]$ in K and for each $B \in \mathcal{T}$), then $\lambda = \lambda_0$.

Proof:

Let A be a fixed set in \mathcal{F} . $\lambda_n(A \times \cdot)$ denotes the restriction of λ_n to A and is a measure on (Y, \mathcal{T}) . It is easy to verify that $\lambda_n(A \times \cdot)$ tends weakly to a measure, $\lambda(A \times \cdot)$ say, on (Y, \mathcal{T}) .

Let C be any compact μ -continuity set in K . By Vitali's covering theorem (Theorem 2.2) there is a sequence of disjoint sets $\{A_n\} \in \mathcal{F}$ such that if $D = C - \bigcup_1^\infty A_n$ then $\mu(D) = 0$. Define for each $B \in \mathcal{T}$, $\lambda(C \times B) = \sum_1^\infty \lambda(A_n \times B)$. It is easy to see that $\lambda(C \times B)$ is well defined for each $C \in \mathcal{C}$, the class of all compact μ -continuity sets in K . X being locally compact separable metric, \mathcal{C} constitutes a base for the Borel field \mathcal{S} and hence $\lambda(C \times B)$ can be uniquely extended as a measure on $\mathcal{S} \times \mathcal{T}$.

Since the sequence of probability measures μ_n converges weakly to μ , given $\varepsilon > 0$ there is a set $C \in \mathcal{C}$ such that $\mu_n(C) > 1 - \varepsilon$ for all n . Again by Vitali's covering theorem (Theorem 2.2) there is a finite collection of disjoint set $A_1, \dots, A_m \in \mathcal{F}$ such that if $D = C - \bigcup_1^m A_r$ then $\mu(D) < \varepsilon$. Hence $\mu_n(D) < 2\varepsilon$ if $n \geq n_0$. Also, since $\nu_n(A_r, \cdot)$ converges weakly for each r , $r = 1, \dots, m$, there is a compact set $H \in \mathcal{T}$ such that

$\nu_n(A_r \times H) > 1 - \varepsilon$ for $r = 1, \dots, m$ and all n . Thus

$\lambda_n(C \times H) \geq 1 - 4\epsilon$ for all $n \geq n_0$. Thus the sequence of measures $\{\lambda_n\}$ is compact under weak convergence.

If λ' is any limit point of $\{\lambda_n\}$, we note that λ and λ' agree on compact continuity sets, and so are identical. Thus $\lambda_n \rightarrow \lambda$.

The second part of the theorem is a well known sufficient condition (X is separable metric) for λ to be equal to λ_0 . (See Doob (1953)).

Before presenting the last of our main theorems we prove several lemmas.

In the following lemmas, I is a compact metric space and $M = R_k$. $g(\theta, m)$ is a bounded continuous function on $I \times M$. J is any bounded interval in M . $\{\nu_n(\theta, \cdot)\}; n = 0, 1, \dots$ is a family of sequences of probability measures on M indexed by θ in I .

$\{\nu_n(\theta, \cdot)\}$ converges to $\nu(\theta, \cdot)$ the UC^* sense relative to $\theta \in I$.
Let

$$v_n(\theta) = \int g(\theta, m) \nu_n(\theta, dm), \quad n = 0, 1, \dots \quad (14)$$

2.3.5. Lemma.

$g(\theta, m)$ is equicontinuous in θ at each m ,

i.e. if $\theta_n \rightarrow \theta_0$, then $g(\theta_n, m) \rightarrow g(\theta_0, m)$ uniformly in $m \in J$.

Proof:

This follows immediately from the uniform continuity of $g(\theta, m)$ on $I \times J$.

2.3.6. Lemma

$$\sup_{\Theta} |v_n(\Theta) - v_0(\Theta)| \rightarrow 0 \quad \text{as } n \rightarrow \infty \quad \dots \quad (15)$$

Proof:

This lemma is a simple corollary of the general results of Ranga Rao (1960, 1961). We however, give here a short proof for the sake of continuity.

If (15) is not true, then there is a sequence $\{\Theta_n\}$ and $\alpha > 0$, such that

$$|v_n(\Theta_n) - v_0(\Theta_n)| \geq \alpha \quad \text{for each } n \dots \quad (16)$$

Since I is compact, there is a subsequence $\{\Theta_{n_r}\}$ such that $\Theta_{n_r} \rightarrow \Theta_0$ as $r \rightarrow \infty$.

We then have

$$\begin{aligned} & |v_{n_r}(\Theta_{n_r}) - v_0(\Theta_{n_r})| \\ &= \left| \int g(\Theta_{n_r}, m) \nu_{n_r}(\Theta_{n_r}, dm) - \int g(\Theta_{n_r}, m) \nu_0(\Theta_{n_r}, dm) \right| \\ &\leq \int |g(\Theta_{n_r}, m) - g(\Theta_0, m)| \nu_{n_r}(\Theta_{n_r}, dm) \\ &+ \left| \int g(\Theta_0, m) \nu_{n_r}(\Theta_{n_r}, dm) - \int g(\Theta_0, m) \nu_0(\Theta_0, dm) \right| \\ &+ \left| \int g(\Theta_0, m) \nu_0(\Theta_0, dm) - \int g(\Theta_0, m) \nu_0(\Theta_{n_r}, dm) \right| \\ &+ \int |g(\Theta_0, m) - g(\Theta_{n_r}, m)| \nu_0(\Theta_{n_r}, dm). \end{aligned}$$

Given any $\epsilon > 0$ the first, second, third and fourth terms on the right hand side can each be made $< \epsilon$ by using lemma 3.5 and (5), (5), (7), and lemma 3.5 and (7), respectively if $r \geq R$. Since $\epsilon > 0$ is arbitrary, this is a contradiction to (16). Hence the lemma.

2.3.7. Lemma

$v_0(\theta)$ is continuous in θ .

Proof:

This follows immediately from (7).

Let $\{u_n(m)\}$ be a sequence of functions on a separable complete metric space M converging to a bounded continuous function $u(m)$ uniformly on every compact set. Let $\{\mu_n\}$ be a sequence of probability measures on M converging weakly to μ . We then have the following lemma.

2.3.8. Lemma

$$\int u_n(m) d\mu_n \rightarrow \int u(m) d\mu.$$

Proof:

The proof is immediate.

In the following theorem $Y = R_k$ and X is any separable complete metric space. Here we impose certain conditions on the conditional distributions similar to those employed by Parzen (1954), Steck (1957) and others.

2.3.9. Theorem.

Let the sequence of marginal distributions μ_n converge weakly to μ . Let the sequence of conditional distributions

$\nu_n(x, \cdot)$ converge to $\nu(x, \cdot)$ in the UC^* sense relative to $x \in I$ for every compact subset I of X . Then the sequence λ_n of joint distributions converges weakly to λ_0 .

Proof:

Let $g(x, y)$ be any bounded continuous function on $X \times Y$.

Let $u_n(x) = \int_Y g(x, y) \nu_n(x, dy)$, $u(x) = \int_Y g(x, y) \nu(x, dy)$.

Lemma 3.6 shows that $u_n(x) \rightarrow u(x)$ uniformly in $x \in I$ for every compact $I \subset X$.

Now,

$$\begin{aligned} \int g(x, y) d\lambda_n &= \int_X d\mu_n \int_Y g(x, y) \nu_n(x, dy) \\ &= \int u_n(x) d\mu_n \rightarrow \int u(x) d\mu \text{ by lemma 3.8} \\ &= \int_X d\mu \int_Y g(x, y) \nu(x, dy) \\ &= \int g(x, y) d\lambda \end{aligned}$$

where $\lambda(C) = \int_X d\mu \int_{C_x} \nu(x, dy)$

with $C_x = \{y : (x, y) \in C\}$. See Halmos (1950).

Thus $\lambda = \lambda_0$ and $\lambda_n \rightarrow \lambda_0$.

2.4. A counterexample.

We present below an example to show that some such conditions, as imposed in theorems 3.2, 3.3, 3.4, 3.9 are, in general necessary.

X and Y are the real line and S and T the usual field of Borel sets. The random variable (ξ_n, η_n) takes the values

$$\left(\frac{1}{n}, \frac{1}{n}\right), \left(\frac{1}{n}, 1 + \frac{1}{n}\right), \left(1 + \frac{1}{n}, \frac{1}{n}\right) \text{ and } \left(1 + \frac{1}{n}, 1 + \frac{1}{n}\right)$$

with probabilities $\frac{1}{8}, \frac{3}{8}, \frac{3}{8},$ and $\frac{1}{8}$ respectively if n is even and with probabilities $\frac{3}{8}, \frac{1}{8}, \frac{1}{8}, \frac{3}{8}$ respectively if n is odd.

It is easy to see that the marginal distributions of ξ_n and η_n converge weakly to the same distribution with masses $\frac{1}{2}$ and $\frac{1}{2}$ at 0 and 1. The conditional distributions are trivially convergent. The joint distributions do not converge.

2.5. The use of theorems 3.2, 3.3, 3.4, 3.9.

First consider the following well known theorem in asymptotic distribution theory:

2.5.1. THEOREM (Varadarajan (1958)).

If \underline{X}_n is a sequence of distributions on R_k , the Euclidean space of k dimensions, then $\underline{X}_n \Rightarrow \underline{X}$ if and only if $L \underline{X}_n \Rightarrow X_L$ for every linear transformation L on R_k .

From this theorem we can deduce elegantly the multivariate central limit theorem from the univariate one, and the joint asymptotic

distribution of several U-statistics (for definition etc., see Hoeffding (1948)) from that of a single U-statistic. All that is required is the trivial verification that any linear combination of the variables in the multivariate case is a variable satisfying the conditions of the central limit theorem for the univariate case.

Theorems 3.2, 3.3, 3.4, 3.9 should prove as useful as the above theorem (Theorem 5.1). For instance, we can deduce the joint asymptotic distribution of several sample quantiles from that of a single quantile. As an illustration, assuming the theorem for the asymptotic distribution of the sample median (see Smirnov (1949), Cramer (1946)), let us show that the sample first quartile and the median are jointly asymptotically normally distributed.

Let z_1, \dots, z_{4n+3} be $4n+3$ independent observations on a random variable Z with distribution function $F(z)$ which possesses a density function $f(z)$. It is assumed that $f(z)$ is continuous and nonzero at θ , the population median and at δ , the population first quartile.

Let us denote $\sqrt{n}(z_{(2n+2)} - \theta)$, $\sqrt{n}(z_{(n+1)} - \delta)$ by (ξ_n, η_n) where $z_{(2n+2)}$ is the sample median and $z_{(n+1)}$ is the sample first quartile. When ξ_n is fixed at x , η_n is the normalised median of a sample of size $(2n+1)$ on the random variable Z truncated to the region $(-\infty, \theta + \frac{x}{\sqrt{n}})$, and hence is asymptotically normal. Some algebraic computations show that the mean and variance of the limiting conditional distribution are $\frac{x f(\theta)}{2 f(\delta)}$ and $\frac{1}{32 f^2(\delta)}$. It is well

known that ξ_n tends strongly to the normal distribution with mean zero and variance $\frac{1}{16 f^2(\theta)}$ (since the densities converge pointwise - Cramer (1946)). An application of theorem 3.3 shows that the joint distribution of (ξ_n, η_n) converges to the bivariate normal distribution with means zero, zero, variances $\frac{1}{16 f^2(\theta)}$, $\frac{3}{64 f^2(\theta)}$ and correlation $\frac{1}{\sqrt{3}}$.

It is to be hoped that the technique illustrated above will be available to establish many results concerning limit distributions. The following chapters, which are mainly applications of the theorems of the present chapter, serve to illustrate and emphasize the usefulness and power of the present technique.

Chapter III

APPLICATIONS TO STATISTICAL METHODS

3.1. Introduction and summary.

The basic tools required in this work were developed in Chapter II. We now proceed to use them in the various fields of Statistics. In the present chapter we deal with the applications in Statistical Methods.

Our theorems being primarily intended to obtain joint asymptotic distributions of several random variables, it is natural to apply them to classes of statistics with known asymptotic marginal distributions but with, as yet, unknown joint asymptotic behaviour. One such class is the class of U-statistics and order statistics.

It is well known that U-statistics have asymptotically normal distributions. See Hoeffding (1948). Though both U-statistics and order statistics are symmetric functions of the observations, their asymptotic behaviours are different. The asymptotic distribution of the sample quantile depends purely on the behaviour of the parent distribution function at the population quantile, and is restricted to just four types of distributions. See Smirnov (1949), (1952). The normal distribution is one of these limiting forms. Similarly the asymptotic distribution of the extreme value depends on the behaviour of the parent distribution function at the extremes, and

is limited to three types of distributions. See Gnedenko (1943) and Smirnov (1949). None of these limiting types is normal. The asymptotic distribution of a third class of order statistics has been considered by Kawata (1951) and has been shown to be normal under certain conditions on the parent distribution function. These important results have been reproduced in some detail in section 2.

We therefore pose a very important and interesting question: Does the joint asymptotic distribution of a U-statistic and an order statistic exist and if so what is it? What precisely are the allowable types of limiting distributions? Using only the known condition necessary for the existence of the marginal asymptotic distributions, we give a complete answer to this question and also determine precisely the only allowable limit types of distributions, in section 4.

As a review of the previous work in this direction, we remark that Sukhatme (1957) investigated this question in the case when the order statistic is the median. A more general situation was studied later by Sethuraman and Sukhatme (1959). We should state that these authors made use of the technique of direct evaluation of characteristic functions and this obscured the inherent simplicity of the problem, and so that they were led to impose various unnecessary conditions. We shall see in section 4 how their results have been made complete and conclusive.

In section 5 we present certain easy and straightforward generalisations to the case of several U-statistics, several order statistics and generalised U-statistics (for definition see Lehmann (1951) and section 2.)

As a further generalisation we establish, in section 6, joint asymptotic distributions of a similar nature for the class of V-statistics and order statistics. V-statistics can be loosely said to be U-statistics depending on a parameter, where the parameter is replaced by some suitable statistic.

As a different type of application of the basic theorems of Chapter II, we provide in section 7 some elegant proofs of the results of Hoeffding (1953) concerning the distributions of mixtures of order statistics. We replace the algebraic approach of Hoeffding (1953) which is not easily amenable to generalisations, by another approach which not only gives some insight into the problem but also permits easy extensions in several directions.

3.2. Notations, definitions and basic preliminaries.

The notations, definitions etc., of this chapter, unless indicated otherwise, will be those developed in this section.

Let X be a random variable on the real line with distribution function $F(x)$. Whenever the density function exists we shall denote it by $f(x)$.

Let x_1, x_2, \dots, x_n be n independent observations on X . Arranging the observations in increasing order of magnitude we get,

$x_{(1)}^{(n)}, x_{(2)}^{(n)}, \dots, x_{(n)}^{(n)}, x_{(r)}^{(n)}$ is called the r th order statistic. The corresponding random variable is denoted by $X_{(r)}^{(n)}$ and the distribution function, by $F(x; r, n)$. From now on we will omit the superscripts among the order statistics and write them simply as $x_{(r)}, X_{(r)}$ etc.

Let $\{a_n\}$ be a sequence of integers with $1 \leq a_n \leq n$.

One of the statistics of primary interest in this chapter is $X_{(a_n)}$. We know that its distribution function, $F(x; a_n, n)$, is given by

$$F(x; a_n, n) = \sum_{r=a_n}^n \binom{n}{r} F^r(x) (1 - F(x))^{n-r} \dots \quad (1)$$

A complete and exhaustive treatment of the asymptotic distribution of $X_{(a_n)}$ for two kinds of sequences $\{a_n\}$ has been given in Gnedenko (1943) and Smirnov (1949), though this topic has already been touched upon by others, e.g. Fisher and Tippett (1928), Erchéet (1927), Gumbel, Cramer etc., in a fragmentary fashion. A third kind of sequence $\{a_n\}$, infrequently met with in practice, was considered and discussed by Kawata (1951) under certain conditions on $F(x)$. We paraphrase below some of the important results from Gnedenko (1943), Smirnov (1949) and Kawata (1951) to suit our requirements.

Case A.

$$\sqrt{n} \left(\frac{a_n}{n} - p \right) \rightarrow 0 \quad \text{as } n \rightarrow \infty, \quad 0 < p < 1 \dots \quad (2)$$

The necessary and sufficient condition that for some $\alpha_n > 0, \beta_n$

$$\xi_n = \frac{X_{(a_n)} - \beta_n}{\alpha_n} \dots \quad (3)$$

tends weakly to a proper* random variable ξ , (i.e. $F(\alpha_n x + \beta_n, a_n, n)$) tends in law to a proper* distribution function $\Phi_A(x)$ is that

$$u_n(x) = (F(\alpha_n x + \beta_n) - p) (n/pq)^{\frac{1}{2}} \rightarrow u(x) \dots \quad (4)$$

as $n \rightarrow \infty$, where $u(x)$ is a non-decreasing function. In this case the limiting distribution $\Phi_A(x)$ is determined by the relation

$$\Phi_A(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u(x)} \exp(-t^2/2) dt \dots \quad (5)$$

Further, $u(x)$ necessarily satisfies the relation: For every integer $m > 1$, there exist constants $k_m > 0$, l_m such that

$$u(x) = \sqrt{m} u(k_m x + l_m) \dots \quad (6)$$

There are only four types of functions $u(x)$ which satisfy (6) (i.e.

there are only four types* of limiting distributions $\Phi_{A_1}(x)$, $\Phi_{A_2}(x)$,

$\Phi_{A_3}(x)$ and $\Phi_{A_4}(x)$.) These four functions together with the correspon-

ding necessary and sufficient conditions A_1, A_2, A_3 and A_4 on $F(x)$

* A random variable ξ is said to be 'proper' if ξ is not a constant with probability one. A distribution function is said to be 'proper' if it is the distribution of a proper random variable.

+ Two distribution functions $F_1(x)$ and $F_2(x)$ are said to belong to the same type if there are constants $a > 0$, b such that

$$F_1(x) = F_2(ax + b).$$

are as follows.

A_1 .

Condition A_1 :

There is a θ such that

$$F(\theta + 0) = p \text{ and } F(\theta + \varepsilon) > p \text{ for each } \varepsilon > 0.$$

$$\frac{F(\theta + x) - p}{p - F(\theta - x)} \rightarrow 0 \text{ as } x \rightarrow +0$$

For each $\tau > 0$,

$$\frac{F(\theta + \tau x) - p}{F(\theta + x) - p} \rightarrow \tau^\alpha \text{ as } x \rightarrow +0 \text{ for some } \alpha > 0.$$

$$u(x) = \begin{cases} -\infty & x < 0 \\ c x^\alpha & x > 0, \quad c > 0, \quad \alpha > 0 \end{cases} \dots \quad (7)$$

A_2 .

Condition A_2 :

$$F(\theta - 0) = p, \quad F(\theta - \varepsilon) < p \text{ for each } \varepsilon > 0$$

$$\frac{p - F(\theta - x)}{F(\theta + x) - p} \rightarrow 0 \text{ as } x \rightarrow 0$$

For each $\tau > 0$

$$\frac{p - F(\theta - \tau x)}{p - F(\theta - x)} \rightarrow \tau^\alpha \text{ as } x \rightarrow +0.$$

$$u(x) = \begin{cases} -c |x|^\alpha & \text{if } x < 0, \quad c > 0, \quad \alpha > 0 \\ \infty & \text{if } x > 0 \end{cases} \dots \quad (8)$$

A₃.

Condition A₃:

There is a θ such that $F(x)$ is continuous at θ and $F(\theta) = p$, $F(\theta - \varepsilon) < p < F(\theta + \varepsilon)$ for each $\varepsilon > 0$.

$\frac{F(\theta + x) - p}{p - F(\theta - x)} \rightarrow A$ as $x \rightarrow +0$, where A is a positive constant.

For each $\tau > 0$

$$\frac{F(\theta + \tau x) - p}{F(\theta + x) - p} \rightarrow \tau^\alpha \quad \text{as } x \rightarrow +0$$

$$u(x) = \begin{cases} -c_1 |x|^\alpha & x < 0, c_1 > 0, \alpha > 0 \\ c_2 x^\alpha & x > 0, c_2 > 0 \end{cases} \quad \dots \quad (9)$$

A₄.

Condition A₄:

Let $\frac{a}{p} = \sup \{x: F(x) < p\}$, $\bar{a}_p = \inf \{x: F(x) > p\}$.

One of the following is satisfied

i) $\frac{a}{p} < \bar{a}_p$

ii) $\frac{a}{p} = \bar{a}_p = a_p$

$F(a_p + 0) = p$, $F(a_p - 0) < p$.

For each $\tau > 0$,

$$\frac{F(a_p + \tau x) - p}{F(a_p + x) - p} \rightarrow +\infty \quad \text{as } x \rightarrow +0$$

iii) $\frac{a}{p} = \bar{a}_p = a_p$

$F(a_p - 0) = p$, $F(a_p + 0) > p$.

For each $\tau > 0$,

$$\frac{p - F(a_p - \tau x)}{p - F(a_p - x)} \rightarrow +\infty \quad \text{as } x \rightarrow +0$$

$$\text{iv) } \frac{a}{p} = \bar{a}_p = a_p$$

$$F(a_p - 0) = F(a_p + 0) = p$$

$$\sqrt{n} \left\{ F\left(a_p + \frac{\alpha_n - \beta_n}{2} + \frac{\alpha_n + \beta_n}{2} x\right) - p \right\} \rightarrow \begin{cases} -\infty & \text{if } x < -1 \\ 0 & \text{if } -1 < x < 1 \\ +\infty & \text{if } x > 1 \end{cases}$$

where $\alpha_n > 0$, $\beta_n > 0$ are determined by the relations

$$\alpha_n = \inf \left\{ x : F(a_p + x(1 - 0)) - p \leq \frac{1}{\sqrt{n}} \leq F(a_p + x(1 + 0)) - p \right\}$$

$$\beta_n = \inf \left\{ x : F(a_p - x(1 + 0)) - p \leq \frac{1}{\sqrt{n}} \leq F(a_p - x(1 - 0)) - p \right\}$$

$$u(x) = \begin{cases} -\infty & x < -1 \\ 0 & -1 < x < 1 \\ +\infty & x > 1 \end{cases} \quad \dots \quad (10)$$

Case B

$$a_n = k, \text{ a constant for all } n \quad \dots \quad (11)$$

We remark at the outset that the case $a_n = n - k$, k a constant, is dealt with analogously.

The necessary and sufficient condition that for suitable constants

$\alpha_n > 0$, β_n

$$\xi_n = \frac{X(a_n) - \beta_n}{\alpha_n} \quad \dots \quad (12)$$

tends weakly to a proper random variable ξ (i.e. $F(\alpha_n x + \beta_n; a_n, n)$) tends in law to a proper distribution function $\Phi_B(x)$ is that

$$v_n(x) = n F(\alpha_n x + \beta_n) \rightarrow v(x) \quad \dots \quad (13)$$

as $n \rightarrow \infty$, where $v(x)$ is a nondecreasing function of x .

In this case, the limiting distribution $\Phi_B(x)$ is determined by the relation

$$\begin{aligned} \Phi_B(x) &= \frac{1}{(k-1)!} \int_0^{v(x)} e^{-t} t^{k-1} dt \\ &= \sum_{r=k}^{\infty} \frac{e^{-v(x)} (v(x))^r}{r!} \quad \dots \quad (14) \end{aligned}$$

Further, $v(x)$ necessarily satisfies the relation: For every integer $m > 1$, there exist constants $k_m > 0$, l_m such that

$$v(x) = n v(k_m x + l_m) \quad \dots \quad (15)$$

There are only three types of functions $v(x)$ satisfying (15), i.e., there are only three types of limit laws $\Phi_{B_1}(x)$, $\Phi_{B_2}(x)$ and $\Phi_{B_3}(x)$. These three functions $v(x)$ together with the corresponding necessary and sufficient conditions B_1 , B_2 and B_3 on $F(x)$ are as follows:

B_1 Condition B_1 There is a $\theta (< \infty)$ such that $F(\theta) = 0, F(\theta + \varepsilon) > 0$ for each $\varepsilon > 0$.For any $\tau > 0$,

$$\frac{F(\theta + \tau x)}{F(\theta + x)} \rightarrow \tau^\alpha \quad \text{as } x \rightarrow +\infty$$

$$v(x) = \begin{cases} 0 & x < 0 \\ x^\alpha & x > 0 \end{cases} \quad \dots \quad (16)$$

 B_2 Condition B_2 For each $\tau > 0$,

$$\frac{F(x)}{F(\tau x)} \rightarrow \tau^\alpha \quad \text{as } x \rightarrow -\infty.$$

$$v(x) = \begin{cases} |x|^{-\alpha} & x < 0 \\ \infty & x > 0 \end{cases} \quad \dots \quad (17)$$

 B_3 Condition B_3

Define

$$\beta_n = \inf \left\{ x: F(x-0) \leq \frac{1}{n} \leq F(x+0) \right\}$$

$$\text{and } \alpha_n = \inf \left\{ x: F(\beta_n - x(1+0)) \leq \frac{1}{n e} \leq F(\beta_n - x(1-0)) \right\}$$

Then $n F(\alpha_n x + \beta_n) \rightarrow e^x$ for each x , as $x \rightarrow \infty$.

$$v(x) = e^x, \quad -\infty < x < \infty \quad \dots \quad (18)$$

Case C.

$$\frac{a_n}{n} \rightarrow 1, \quad n - a_n \rightarrow \infty \quad \text{as } n \rightarrow \infty \quad \dots \quad (19)$$

This case has not been completely investigated in the literature. The best known result in this direction is due to Kawata (1951) and is as follows.

Condition C.

$F(x)$ satisfies the relation

$$\frac{1 - F(x(1 + \varepsilon(x)))}{1 - F(x)} = 1 - \alpha \varepsilon(x) + o(\varepsilon(x)) \quad \dots \quad (20)$$

as $x \rightarrow \infty$, where $\varepsilon(x)$ is a function of x which tends to zero as $x \rightarrow \infty$ and $\alpha > 0$ is a fixed constant.

Define

$$\beta_n = \inf \left\{ x: F(x - 0) \leq \frac{a_n}{n} \leq F(x + 0) \right\} \dots \quad (21)$$

$$\alpha_n = \frac{\beta_n}{\alpha} \sqrt{\frac{a_n}{n(n - a_n)}} \quad \dots \quad (22)$$

Then it can be shown that

$$\frac{F(\alpha_n x + \beta_n) - \frac{a_n}{n}}{\sqrt{\frac{a_n(n - a_n)}{n^3}}} \rightarrow x \quad \text{as } n \rightarrow \infty \quad \dots \quad (23)$$

In this case if

$$\xi_n = \frac{X_{(a_n)} - \beta_n}{\alpha_n} \quad \dots \quad (24)$$

then ξ_n converges weakly to a proper random variable ξ with distribution function $\int_G(x)$ given by

$$\phi_C(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-t^2/2) dt \quad \dots \quad (25)$$

We now proceed to recapitulate the definition and asymptotic properties of a U-statistic. The term U-statistic is due to Hoeffding (1948), but we use a slightly different definition.

Let $\psi(x_1, \dots, x_s)$ be a function symmetric in its arguments and possess a finite second moment; i.e.

$$E(\psi^2(X_1, \dots, X_s))^* < \infty \quad \dots \quad (26)$$

Then a U-statistic, U_n based on a sample of size n is defined as

$$U_n = \frac{1}{\binom{n}{s}} \sum \psi(x_{i_1}, x_{i_2}, \dots, x_{i_s}) \quad \dots \quad (27)$$

where the summation is made over all combinations (i_1, i_2, \dots, i_s) of s indices out of $(1, \dots, n)$. Here $\psi(x_1, \dots, x_s)$ is said to be the kernel of the U-statistic U_n .

Define

$$\varphi(x) = E \left\{ \psi(x_1, x_2, \dots, x_s) \right\}^* \quad \dots \quad (28)$$

We give below the main result of Hoeffding (1948) which we shall use in this chapter.

$$\text{Let } E(\varphi(X)) = m \quad \text{and} \quad V(\varphi(X)) = \sigma^2 \quad \dots \quad (29)$$

* By $E \xi$ and $V \xi$ we mean the expectation and variance, respectively, of a random variable ξ .

Let

$$\eta'_n = \frac{\sum_1^n (\varphi(X_i) - m)}{\sqrt{n}} \quad \dots \quad (30)$$

$$\eta_n = \sqrt{n} (U_n - m) \quad \dots \quad (31)$$

Then η_n and $s \eta'_n$ are asymptotically equivalent, i.e.

$$\eta_n - s \eta'_n \rightarrow 0 \text{ in probability } \dots \quad (32)$$

Thus η_n tends weakly to a random variable η with a normal distribution with mean zero and variance $s^2 \sigma^2$.

Extensions, in several directions, of the above result are known. For instance it is well known that joint asymptotic distribution of several U-statistics is multivariate normal. See Hoeffding (1948) ^{the} and remark in section 2.5.

Another sort of generalisation is due to Lehmann (1951). Here the U-statistic is defined on two independent samples, x_1, \dots, x_{n_1} , y_1, \dots, y_{n_2} , of sizes n_1 and n_2 on two random variables X and Y . Let $\psi(x_1, x_2, \dots, x_{s_1}; y_1, y_2, \dots, y_{s_2})$ be a function symmetric in the x -arguments and the y -arguments separately and possessing a finite second moment, i.e. $E(\psi^2(X_1, X_2, \dots, X_{s_1}; Y_1, Y_2, \dots, Y_{s_2})) < \infty \dots$ (33)

The generalised U-statistic, U_{n_1, n_2} , with kernel $\psi(x_1, x_2, \dots, x_{s_1}; y_1, y_2, \dots, y_{s_2})$, is defined as

$$U_{n_1, n_2} = \frac{1}{\binom{n_1}{s_1} \binom{n_2}{s_2}} \sum \psi(x_{i_1}, x_{i_2}, \dots, x_{i_{s_1}}; y_{j_1}, y_{j_2}, \dots, y_{j_{s_2}}) \dots (34)$$

where the sum is taken over all combinations $(i_1, i_2, \dots, i_{s_1})$ of s_1 x indices from $(1, \dots, n_1)$ and over all combinations $(j_1, j_2, \dots, j_{s_2})$ of s_2 y -indices from $(1, \dots, n_2)$.

Let us define

$$\varphi_1(x_1) = E(\psi(x_1, X_2, \dots, X_{s_1}; Y_1, Y_2, \dots, Y_{s_2})) \dots (35)$$

$$\varphi_2(y_1) = E(\psi(X_1, X_2, \dots, X_{s_1}; y_1, Y_2, \dots, Y_{s_2})) \dots (36)$$

$$E(\varphi_1(X)) = E(\varphi_2(Y)) = m \dots (37)$$

$$V(\varphi_1(X)) = \sigma_1^2 \dots (38)$$

$$V(\varphi_2(Y)) = \sigma_2^2 \dots (39)$$

A result due to Lehmann (1951) (see also Fraser (1957)) says that if

$$\eta_n = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} (U_{n_1, n_2} - m) \dots (40)$$

$$\eta'_n = s_1 \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \left(\frac{1}{n_1} \sum \varphi_1(X_i) - m \right) + s_2 \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \left(\frac{1}{n_2} \sum \varphi_2(Y_j) - m \right)$$

and $\frac{n_1}{n_2} \rightarrow c$, $0 < c < \infty$, then η_n and η'_n are asymptotically equivalent,

$$\text{i.e. } \eta_n - \eta'_n \rightarrow 0 \text{ in probability } \dots (41)$$

so that η_n tends weakly to a random variable η with mean zero and variance $\frac{s_1^2 \sigma_1^2}{1+c} + \frac{c s_2^2 \sigma_2^2}{1+c}$.

In this chapter we introduce yet another generalisation of the U-statistic, which we name as the V-statistic. As a first step we introduce a parameter in the kernel of the U-statistic and later on substitute some sample statistic, usually an order statistic, for the parameter. It is next postulated that the kernel should have certain regularity properties and that the statistic substituted for the parameter should possess a suitable asymptotic behaviour. Then we call this statistic a V-statistic. The details of this case are given in section 6. Such statistics do not seem to have been widely considered in literature. We may in this connection make a reference to Sukhatme (1958) in which V-statistics have been considered, but under conditions more restrictive than ours.

3.3. Auxiliary Theorems.

In this section we state and prove some auxiliary theorems that prove very useful in the verification of UC* convergence of some random variables in section 4.

At the outset we recall a famous theorem due to Gnedenko and Kolmogorov (1954), which we later extend to suit our purposes.

Let

ξ_{11}	-----	ξ_{1k_1}
ξ_{21}	-----	ξ_{2k_2}
-----	-----	-----
ξ_{n1}	-----	ξ_{nk_n}
-----	-----	-----

be a triangular scheme of random variables, in which the random variables in each row are independent.

Let all the random variables entering, have finite second moments.

Let

$$S_n = \frac{\xi_{n1} + \dots + \xi_{nk_n} - \sum_{k=1}^{k_n} E \xi_{nk}}{\sqrt{\sum_{k=1}^{k_n} V \xi_{nk}}} \dots \quad (42)$$

3.3.1. Theorem (Gnedenko and Kolmogorov).

The necessary and sufficient condition that the distribution function of S_n converges in law to the normal distribution function

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-t^2/2) dt \text{ is that, for each } \epsilon > 0,$$

$$\frac{1}{\sum_{k=1}^{k_n} V \xi_{nk}} \int_{|x - E \xi_{nk}| \geq \epsilon \sqrt{\sum_{k=1}^{k_n} V \xi_{nk}}} (x - E \xi_{nk})^2 dF_{nk}(x) \rightarrow 0 \text{ as } n \rightarrow \infty \dots (43)$$

where $F_{nk}(x)$ is the distribution function of ξ_{nk} .

The generalisation that we require is as follows:

Let

$$\begin{matrix} \xi_{11}^\theta, & \dots & \dots & \dots & \xi_{1k_1}^\theta \\ \xi_{21}^\theta, & \dots & \dots & \dots & \xi_{2k_2}^\theta \\ & \dots & \dots & \dots & \\ \xi_{n1}^\theta, & \dots & \dots & \dots & \xi_{nk_n}^\theta \\ & \dots & \dots & \dots & \end{matrix}$$

be a triangular scheme of random variables indexed by Θ , wherein the random variables in each row are independent.

Let all the random variables under consideration have finite second moments.

Let

$$S_n^\Theta = \frac{\xi_{n1}^\Theta + \dots + \xi_{nk_n}^\Theta - \sum_{l=1}^{k_n} E \xi_{nl}^\Theta}{\sqrt{\sum_{l=1}^{k_n} V \xi_{nl}^\Theta}} \dots \quad (44)$$

3.3.2. Theorem.

The necessary and sufficient condition that the distribution function of S_n^Θ converges in law to the normal distribution function

$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-t^2/2) dx$ in the UC* sense relative to Θ is that, for each $\varepsilon > 0$,

$$\frac{1}{\sum_{l=1}^{k_n} V \xi_{nl}^\Theta} \int_{|x - E \xi_{nl}^\Theta| \geq \varepsilon \sqrt{\sum_{l=1}^{k_n} V \xi_{nl}^\Theta}} (x - E \xi_{nl}^\Theta)^2 dF_{nk}^\Theta(x) \rightarrow 0 \dots \quad (45)$$

uniformly in Θ , where F_{nk}^Θ is the distribution function of ξ_{nk}^Θ .

Proof:

This theorem is plain from the definition of UC* convergence in section 2.2 and Gnedenko-Kolmogorov's analysis of theorem 3.1.

Let

$$\begin{array}{r} \xi_{1,1}, \dots, \xi_{1,k(1,\theta)} \\ \xi_{2,1}, \dots, \xi_{2,k(2,\theta)} \\ \dots \\ \xi_{n,1}, \dots, \xi_{n,k(n,\theta)} \\ \dots \end{array}$$

be a triangular scheme of random variables identically and independently distributed in each row. Let the σ_n^2 be the variance in the n th row and there exist quantities c, C such that

$$c < \sigma_n^2 < C \quad \text{for all } n \quad \dots \quad (46)$$

Let $k(n, \theta)$ tend to infinity uniformly in θ , i.e.

$$\inf_{\theta} k(n, \theta) \rightarrow \infty \quad \text{as } n \rightarrow \infty \quad \dots \quad (47)$$

Let

$$S_n^\theta = \frac{\xi_{n1} + \dots + \xi_{n,k(n,\theta)} - k(n, \theta) E \xi_{n1}}{\sqrt{k(n, \theta) V \xi_{n1}}}$$

3.3.3. Theorem.

The distribution of S_n^θ tends in law to the normal distribution function $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-t^2/2) dt$ in the UC^* sense relative to θ

Proof:

This theorem follows immediately from theorem 3.2.

3.4. Asymptotic distribution of U-statistics and order statistics.

In this section we are establishing the asymptotic distributions of U-statistics and order statistics. More precisely, we evaluate the limit laws of (ξ_n, η_n) where ξ_n is the normalised order statistic (say as in (3), (12) or (24)) and η_n is the normalised U-statistic (as in (31)).

Case A.

3.4.1. Theorem.

Let $\{a_n\}$ be a sequence of integers satisfying (2). Let

$$\xi_n = \frac{X_{(a_n)} - \beta_n}{\alpha_n} \quad \text{be as in (3), and } \eta_n = \sqrt{n} (U_n - m) \text{ as in (31).}$$

In order that the distribution function $G_n(x, y)$ of (ξ_n, η_n) converges in law to a proper distribution function $G_A(x, y)$ it is necessary and sufficient that the distribution function of ξ_n converges in law to a proper distribution function, i.e. $F(x)$ satisfies one of the conditions A_1, A_2, A_3 and A_4 .

The limit laws $G_A(x, y)$ are limited to the following four types:

$$G_{A_1}(x, y) = \begin{cases} 0 & \text{if } x < 0 \\ \Lambda(0, x^\alpha, y) & \text{if } x > 0 \end{cases} \quad \dots \quad (48)$$

$$G_{A_2}(x, y) = \begin{cases} \Lambda(-c_1 |x|^\alpha, y) & \text{if } x < 0 \\ \Lambda(\infty, y) & \text{if } x > 0 \end{cases} \dots (49)$$

$$G_{A_3}(x, y) = \begin{cases} \Lambda(-c_1 |x|^\alpha, y) & \text{if } x < 0 \\ \Lambda(c_2 x^\alpha, y) & \text{if } x > 0 \end{cases} \dots (50)$$

$$G_{A_4}(x, y) = \begin{cases} 0 & \text{if } x < -1 \\ \Lambda(0, y) & \text{if } -1 < x < 1 \\ \Lambda(\infty, y) & \text{if } x > 1 \end{cases} \dots (51)$$

where

$$\Lambda(x, y) = \frac{1}{2\pi} \int_{-\infty}^x \exp(-t^2/2) dt \int_{-\infty}^{Q(y,t)} \exp(-u^2/2) du \dots (52)$$

with

$$Q(y, t) = \frac{y - t s(\nu(\theta) - \mu(\theta)) \sqrt{pq}}{\sqrt{p\sigma^2(\theta) + q} \tau^2(\theta)} \dots (53)$$

(For the definition of $\mu(\theta)$, $\nu(\theta)$, $\sigma(\theta)$ and $\tau(\theta)$ see (56), (57), (58) and (59).)

The necessary and sufficient condition that the limiting distribution is of type G_{A_r} is that $F(x)$ satisfies condition A_r , $r = 1, 2, 3, 4$.

Proof:

It is obvious that the conditions stated in the theorem are necessary. We shall now prove that they are sufficient.

In (32) we have noted that η_n and η'_n (defined in (30)) are asymptotically equivalent. From Cramer's theorem (See Cramer (1946) pp. 254) it is plain that we need only derive the limit laws of (ξ_n, η'_n) .

Let $\{\theta_n\}$ be a sequence of points and $p_n = F(\theta_n)$ tend to p , $0 < p < 1$ and θ be any point such that $F(\theta + 0)$ or $F(\theta - 0) = p$.

Let k_n be the number of observations in our sample that are less than or equal to θ_n . k_n has a binomial distribution with parameters n and p_n . Also if

$$\zeta_n = \frac{k_n - np_n}{\sqrt{np_n q_n}} \quad \dots \quad (54)$$

then the distribution function of ζ_n converges to distribution function of ζ given by

$$P(\zeta \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-t^2/2) dt \quad \dots \quad (55)$$

We now prove a lemma.

3.4.2. Lemma:

The limiting distribution of (η'_n, ζ_n) is a bivariate normal distribution.

Proof:

We will first evaluate the conditional distribution of η'_n given $\zeta_n = z$. The conditional distribution of η'_n given $\zeta_n = z$

is not well defined if $np_n + z \sqrt{np_n q_n}$ is not an integer. We are interested in determining the limiting conditional distribution function of η'_n given $\zeta_n = z$ as a continuous function of z .

Thus to remove the indeterminacy just mentioned we substitute the conditional distribution of η'_n given that there are exactly $m_n = [np_n + z \sqrt{np_n q_n}]^*$ observations less than or equal to θ_n for the conditional distribution of η'_n given $\zeta_n = z$. Our version of the conditional distribution is allowable since

$$P \left\{ [np_n + \zeta_n \sqrt{np_n q_n}] = np_n + \zeta_n \sqrt{np_n q_n} \right\} = 1 \quad \dots (55)$$

Let us adopt the following notations in future.

$$\mu(x) = E(\varphi(X) \mid X \leq x) \quad \dots (56)$$

$$\nu(x) = E(\varphi(X) \mid X > x) \quad \dots (57)$$

$$\sigma^2(x) = V(\varphi(X) \mid X \leq x) \quad \dots (58)$$

$$\tau^2(x) = V(\varphi(X) \mid X > x) \quad \dots (59)$$

Let $Y^{(n)}$ be the random variable X truncated to the region $(-\infty, \theta_n]$ and $Z^{(n)}$ be the random variable X truncated to the region (θ_n, ∞) .

Define

$$\eta_n^{(1)}(\zeta_n) = \frac{\sum_1^{m_n} \varphi(Y_i^{(n)}) - m_n \mu(\theta_n)}{\sqrt{m_n} \sigma(\theta_n)} \quad \dots (60)$$

*[x] denotes the integral part of x .

$$\eta_n^{(2)}(\zeta_n) = \frac{\sum_1^{n-m_n} \varphi(z_j^{(n)}) - (n-m_n) \nu(\theta_n)}{\sqrt{n-m_n} \tau(\theta_n)} \dots \quad (61)$$

where $Y_1^{(n)}, \dots, Y_{m_n}^{(n)}$ are identically and independently distributed like $Y^{(n)}$ and $Z_1^{(n)}, \dots, Z_{n-m_n}^{(n)}$ are identically and independently distributed like $Z^{(n)}$.

Let I be any bounded interval on the real line. We note that $m_n \rightarrow \infty$ uniformly for $z \in I$. Also $\sigma(\theta_n)$ are bounded and bounded away from zero. Thus from Theorem 4.3 it follows that the distribution functions of $\eta_n^{(1)}(\zeta_n)$ given $\zeta_n = z$ and $\eta_n^{(2)}(\zeta_n)$ given $\zeta_n = z$ converge to the normal distribution function

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-t^2/2) dt \quad \text{in the UC* sense relative to } z \in I. \quad \text{We}$$

have the relation:

$$\eta_n' = \sqrt{\frac{m_n}{n}} \sigma(\theta_n) \eta_n^{(1)}(\zeta_n) + \sqrt{\frac{n-m_n}{n}} \tau(\theta_n) \eta_n^{(2)}(\zeta_n) \\ - \frac{\frac{m_n}{n} \mu(\theta_n) + (n-m_n) \nu(\theta_n) - n\mu(\infty)}{\sqrt{n}} \dots \quad (62)$$

It can be verified that

$$\sqrt{\frac{m_n}{n}} \sigma(\theta_n) \rightarrow \sqrt{p} \sigma(\theta), \quad \sqrt{\frac{n-m_n}{n}} \tau(\theta_n) \rightarrow \sqrt{q} \tau(\theta) \\ \text{and } \frac{\frac{m_n}{n} \mu(\theta_n) + (n-m_n) \nu(\theta_n) - n\mu(\infty)}{\sqrt{n}} \rightarrow z \sqrt{pq} (\mu(\theta) - \nu(\theta))$$

uniformly in $z \in I \quad \dots$

(63)

From (62) and (63) it follows that the distribution function of η'_n given $\zeta_n = z$ converges $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{R(y,z)} \exp(-t^2/2) dt$ in the UC* sense relative to $z \in I$ where

$$R(y,z) = (y+z \sqrt{pq}(\mu(\theta) - \nu(\theta))) / \sqrt{p\sigma^2(\theta) + q\tau^2(\theta)} \dots \quad (64)$$

We note that this limiting conditional distribution is continuous in z . The lemma now follows from theorem 2.3.9.

Now let $F(x)$ satisfy condition A_1 . Then the limit law of ξ_n is $\Phi_{A_1}(x)$. Thus $P\{\xi_n < x\} \rightarrow 0$ when $x < 0$. Hence

$$P\{\xi_n < x, \eta'_n < y\} \rightarrow 0 \text{ if } x < 0 \dots \quad (65)$$

Now let $x > 0$. Put $\theta_n = \alpha_n x + \beta_n$ in the above lemma (lemma 4.2).

The following events are equivalent

$$\begin{aligned} & \{ \xi_n < x \} ; \{ X_{(a_n)} \leq \alpha_n x + \beta_n \} ; \{ k_n \geq a_n \} ; \\ \text{and } & \left\{ \zeta_n \geq \frac{a_n - np_n}{\sqrt{np_n q_n}} \right\} ; \left\{ \zeta_n \geq -c x^\alpha + o(1) \right\} . \end{aligned}$$

Now, since the limiting distribution function of (η'_n, ζ_n) is continuous, the probabilities converge uniformly over rectangle sets (Polya's theorem). Thus

$$P \{ \xi_n \leq x, \eta'_n < y \} = P \{ \eta'_n \leq y, \zeta_n \geq -cx^\alpha + o(1) \}$$

$$\rightarrow P \{ \eta'_n \leq y, -\zeta_n \leq cx^\alpha \} \dots \quad (66)$$

From (65) and (66) we see that the limiting distribution of (ξ_n, η_n) is $G_{A_1}(x, y)$.

The theorem is completely proved by establishing similar results separately under each of the assumptions that $F(x)$ satisfies conditions A_2, A_3 and A_4 .

Case B.

3.4.3. Theorem.

Let $\{a_n\}$ be a sequence of integers satisfying (11).

Let

$$\xi_n = \frac{X_{(a_n)} - \beta_n}{\alpha_n} \text{ as in (12), and } \eta_n = \sqrt{n} (U_n - m) \text{ as in (31).}$$

In order that the distribution function $G_n(x, y)$ of (ξ_n, η_n) converges in law to a proper distribution function $G_B(x, y)$ it is necessary and sufficient that the distribution function of ξ_n converges in law to a proper distribution function, i.e. $F(x)$ satisfies one of the conditions B_1, B_2 and B_3 .

The limit laws $G_B(x, y)$ are limited to the following three types:

$$G_{B_1}(x,y) = \begin{cases} 0 & \text{if } x < 0 \\ \sum_{r=k}^{\infty} e^{-x^r} \frac{x^{rx}}{r!} x \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y/\Delta\sigma} \exp(-t^2/2) dt & \text{if } x > 0 \end{cases} \dots (67)$$

$$G_{B_2}(x,y) = \begin{cases} \sum_{r=k}^{\infty} e^{-|x|^r} \frac{|x|^{-rd}}{r!} x \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y/\Delta\sigma} \exp(-t^2/2) dt & \text{if } x < 0 \\ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y/\Delta\sigma} \exp(-t^2/2) dt & \text{if } x > 0 \end{cases} \dots (68)$$

$$G_{B_3}(x,y) = \sum_{r=k}^{\infty} e^{-x^r} \frac{x^{rx}}{r!} x \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y/\Delta\sigma} \exp(-t^2/2) dt \quad -\infty < x < \infty$$

... (69)

The necessary and sufficient condition that the limiting distribution is of type G_{B_r} is that $F(x)$ satisfies condition B_r ; $r = 1, 2, 3$.

Proof:

The proof of this theorem is on the same lines as that of theorem 4.1. We only give the statement of a lemma, analogous to lemma 4.2, which is required in the proof.

Let $\{\theta_n\}$ be a sequence of points and $p_n = F(\theta_n)$ tend to zero in such a way that np_n tends to a constant c . We then know that if $\zeta_n = np_n$ then the random variable ζ_n converges strongly to a random variable ζ with a Poisson distribution with parameter c .

3.4.4. Lemma:

The distribution function of (η'_n, ξ_n) converges to the following distribution function

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x/y/\sigma} \exp(-t^2/2) dt \times \sum_{0 \leq r \leq s} \frac{e^{-y/\sigma}}{r!} \dots \quad (70)$$

Proof:

The proof of this lemma is similar to that of lemma 4.2.

Case C.

3.4.5. Theorem.

Let $\{a_n\}$ be a sequence of integers satisfying (19).

Let $\xi_n = \frac{X(a_n) - \beta_n}{\alpha_n}$ as in (24), and $\eta_n = \sqrt{n}(U_n - m)$ as in (31).

Suppose $F(x)$ satisfies condition C. Then the distribution function $G_n(x, y)$ of (ξ_n, η_n) converges in law to $G_C(x, y)$, where

$$G_C(x, y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-t^2/2) dt \times \int_{-\infty}^{y/\sigma} \exp(-u^2/2) du \dots \quad (71)$$

Proof:

We state without proof a lemma analogous to lemmas 4.2 and 4.4.

Let $\{\theta_n\}$ be a sequence of points such that $p_n = F(\theta_n)$ tends to 1 and $nq_n \rightarrow \infty$ as $n \rightarrow \infty$. Let k_n be the number of observations in our sample less than or equal to θ_n . Let

$$\zeta_n = \frac{k_n - np_n}{\sqrt{np_n q_n}} \quad \dots \quad (72)$$

We can easily see that the distribution function of ζ_n converges in law to $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-t^2/2) dt$.

3.4.6. Lemma:

The distribution function of (η'_n, ζ_n) converges to the distribution function

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y/\alpha_n} \exp(-t^2/2) dt \times \int_{-\infty}^x \exp(-u^2/2) du \quad \dots \quad (73)$$

The proof of this theorem now follows in the same way as in theorem 4.1

3.5. Corollaries of theorems in section 4.

The well known classical theorem for order statistics of the case A is the one given in Cramer (1946) pp. 366, which runs as follows: If $F(x)$ possesses a density function $f(x)$ which is non zero at θ and is continuous in a neighbourhood of θ , where θ is the population p th quantile, then $\zeta_n = \frac{X_{(a_n)} - \beta_n}{\alpha_n}$, with $\beta_n = \theta$,

$\alpha_n = \frac{1}{\sqrt{n} f(\theta)}$ has a limiting normal distribution with mean zero and variance pq . If the population were to satisfy this condition, it would follow from theorem 4.1 that the distribution function of (ξ_n, η_n) converges in law to the bivariate normal distribution function

$$\frac{1}{2\pi} \int_{-\infty}^{\frac{x}{\sqrt{pq}}} \exp(-t^2/2) \int_{-\infty}^{Q(y,t)} \exp(-u^2/2) du \dots \quad (74)$$

$$\text{where } Q(y, t) = \left\{ y - s \sqrt{pq} (\gamma(\theta) - \mu(\theta)) \right\} / \sqrt{pq^2(\theta) + q\tau^2(\theta)} \quad (75)$$

as in (53). This is the type of theorem proved in Sethuraman and Sukhatme (1959), but even here these authors had to impose certain restrictions on the kernel of the U-statistic.

We state without proof the following generalisations of the result of Sethuraman and Sukhatme which are immediate consequences of theorems 4.1, 4.3 and 4.5.

3.5.1. Corollary.

Let $\eta_n^{(1)}, \dots, \eta_n^{(k)}$ be k normalised U-statistics with kernels $\psi_1(x_1, \dots, x_{s_1}), \dots, \psi_k(x_1, \dots, x_{s_k})$ respectively.

Let $\xi_n = \frac{X_{(a_n)} - \beta_n}{\alpha_n}$ be the normalised a_n th order statistic.

Then the distribution function of $(\xi_n, \eta_n^{(1)}, \dots, \eta_n^{(k)})$ converges in law to a proper distribution function if and only if the distribution function of ξ_n converges in law to a proper distribution

function. If the order statistic $X_{(a_n)}$ is of the case A then the limit distribution is restricted to just four types; if the order statistic is of the case B then the limit distribution is restricted to just three types; if the order statistic is of case C and $F(x)$ satisfies condition C then the limit distribution is multivariate normal.

The exact limit distribution and the corresponding necessary and sufficient conditions for the various cases are easily written down.

The limit distribution of two normalised order statistics

$$\xi_n^{(1)} = \frac{X_{(a_n)}^{(1)} - \beta_n^{(1)}}{\alpha_n^{(1)}} \quad \text{and} \quad \xi_n^{(2)} = \frac{X_{(a_n)}^{(2)} - \beta_n^{(2)}}{\alpha_n^{(2)}} \quad \text{can be easily determined}$$

if they belong to the cases A, B and C.

3.5.2. Corollary:

The distribution function of $(\xi_n^{(1)}, \dots, \xi_n^{(r)}, \eta_n^{(1)}, \dots, \eta_n^{(k)})$ converges in law to a proper distribution function if and only if the distribution function of $(\xi_n^{(1)}, \dots, \xi_n^{(r)})$ converges in law to a proper distribution function.

The only possible limit types and the corresponding necessary conditions are obvious if the order statistics belong to the cases A, B and C.

3.5.3. Corollary:

Let η_n be derived from a generalised U-statistic based

on two independent samples of sizes n_1 and n_2 , as in (40).

Let $\xi_n^{(1)}$ be a normalised order statistic from the first sample and $\xi_n^{(2)}$ from the second sample. Let $n_1, n_2 \rightarrow \infty$ such that

$\frac{n_1}{n_2} \rightarrow C, 0 < C < \infty$. Then the distribution function of $(\xi_n^{(1)}, \xi_n^{(2)}, \eta_n)$ converges in law to a proper distribution function if and only if the distribution function of $(\xi_n^{(1)}, \xi_n^{(2)})$ converges in law to a proper distribution function.

3.6. Asymptotic distribution of V-statistics and order statistics.

In this section we introduce the class of V-statistics, which includes the class of U-statistics. We later on prove the important theorem that a V-statistic is asymptotically equivalent to some U-statistic. This theorem yields, as immediate consequences, the limit distribution of a single V-statistic, the limit joint distribution of several V-statistics, and the limit joint distribution of V-statistics and order statistics.

For each t , let $\psi(x_1, \dots, x_s, t)$ be a symmetric function in x_1, \dots, x_s . Let $\psi(x_1, \dots, x_s, t)$ have a finite second moment for each t . Let

$$U_n(t) = \frac{1}{\binom{n}{s}} \sum \psi(x_{i_1}, \dots, x_{i_s}, t) \quad \dots \quad (76)$$

where the summation is done in the usual way. $U_n(t)$ is a

U-statistic for each t .

$$\text{Let } m(t) = E(\psi(X_1, \dots, X_g, t)) \dots \quad (77)$$

$$\eta_n(t) = \sqrt{n} (U_n(t) - m(t)) \dots \quad (78)$$

Let $\psi(x_1, \dots, x_g, t)$ possess the property.

For any t, t' in a fixed neighbourhood $(-k, k)$ of 0 ,

$$|\psi(x_1, \dots, x_g, t) - \psi(x_1, \dots, x_g, t')| \leq A(x_1, \dots, x_g) |t - t'|$$

... (79)

$$\text{and } E(A^2(X_1, \dots, X_g)) < \infty .$$

Let $\zeta_n(X_1, \dots, X_n)$ be a statistic such that given any $\epsilon > 0$, there is a constant Z_ϵ such that

$$P \{ |\zeta_n| \leq Z_\epsilon \} \geq 1 - \epsilon \quad \text{for all } n \dots \quad (80)$$

Let $\{d_n\}$ be a sequence such that

$$n d_n^6 \rightarrow 0 \quad \text{as } n \rightarrow \infty \dots \quad (81)$$

Let

$$V_n = U_n(d_n \zeta_n) \dots \quad (82)$$

Then, if (79), (80) and (81) are satisfied, we call V_n a V-statistic of the first kind.

Now let $\psi(x_1, \dots, x_g, t)$ possess, in stead, the following

properties:

For all t in $(-k, k)$, a fixed neighbourhood of 0 ,

$$| \psi(x_1, \dots, x_s, t) | \leq A \quad \dots \quad (83)$$

where A is a constant.

For all t, t' in $(-k, k)$

$$E | \psi(X_1, \dots, X_s, t) - \psi(X_1, \dots, X_s, t') | \leq K |t - t'| \dots (84)$$

where K is a constant.

Also let $\{e_n\}$ be a sequence such that

$$n e_n^2 \rightarrow \text{a constant as } n \rightarrow \infty \quad \dots \quad (85)$$

$$\text{Let } V_n = U_n(e_n \zeta_n) \quad \dots \quad (86)$$

Then, if (83), (84), (80) and (85) are satisfied, we call V_n a V -statistic of the second kind.

3.6.1. Theorem.

Let V_n be a V -statistic of the first kind. Let

$$\eta_n(d_n \zeta_n) = \sqrt{n} (V_n - m(d_n \zeta_n)) \quad \dots \quad (87)$$

$$\text{Then } \eta_n(d_n \zeta_n) - \eta_n(0) \rightarrow 0 \text{ in probability } \quad \dots \quad (88)$$

Proof:

We first prove the following:

For any constant $Z > 0$,

$$M_n(Z) = \sup_{|z| \leq Z} |\eta_n(d_n z) - \eta_n(0)| \rightarrow 0 \text{ in probability.}$$

Divide the interval $(-Z, Z)$ into $k_n = \left[\frac{2Z}{h_n} \right] + 1$ equal intervals of length h_n , where h_n is chosen as in (97).

We can write

$$\eta_n(d_n z) - \eta_n(0) = \eta_n(d_n z) - \eta_n(d_n r h_n) + \eta_n(d_n r h_n) - \eta_n(0) \dots (89)$$

where r is chosen so that $(r-1)h_n \leq z \leq r h_n$.

From (89) we find

$$M_n(Z) \leq L_n^{(1)}(Z) + L_n^{(2)}(k_n) \dots (90)$$

where

$$L_n^{(1)}(Z) = \sup_{-k_n \leq r \leq k_n} \sup_{(r-1)h_n \leq z \leq r h_n} |\eta_n(d_n z) - \eta_n(d_n r h_n)| \dots (91)$$

and

$$L_n^{(2)}(k_n) = \sup_{-k_n \leq r \leq k_n} |\eta_n(d_n r h_n) - \eta_n(0)| \dots (92)$$

If can be easily verified that

$$E \left\{ \sup_{(r-1)h_n \leq z \leq rh_n} |\eta_n(d_n z) - \eta_n(d_n r h_n)| \right\} \leq \sqrt{n} d_n h_n C_1 \dots (93)$$

where C_1 is a constant.

Thus given any $\theta > 0$, we have by Tchebycheff's lemma

$$P \left\{ L_n^{(1)}(z) > \frac{\theta}{2} \right\} < \frac{2C_1}{\theta} \sqrt{n} d_n h_n \dots (94)$$

We have

$$E(\eta_n(d_n r h_n) - \eta_n(0))^2 \leq d_n^2 r^2 h_n^2 C_2 \dots (95)$$

where C_2 is a constant.

Thus

$L_n^{(2)}(k_n) \rightarrow 0$ in probability if

$$C_2 \times \sum_{r=-k_n}^{k_n} d_n^2 r^2 h_n^2 = C_3 d_n^2 / h_n \rightarrow 0 \text{ as } n \rightarrow \infty \dots (96)$$

Given $\varepsilon > 0$, we choose h_n to satisfy the relation

$$\frac{2C_1}{\theta} \sqrt{n} d_n h_n = \varepsilon \dots (97)$$

and we find using (94), (96), (97) and (81) that

$$P \left\{ M_n(z) > \theta \right\} < 2\varepsilon \quad \text{if } n \geq N(\theta, \varepsilon) \dots (98)$$

The theorem follows from (98) and (80).

3.6.2. Theorem

Let V_n be a V-statistic of the second kind.

$$\text{Let } \eta_n(\theta_n \zeta_n) = \sqrt{n} (V_n - \mathfrak{m}(\theta_n \zeta_n)) \dots \quad (99)$$

Then

$$\eta_n(\theta_n \zeta_n) - \eta_n(0) \rightarrow 0 \text{ in probability } \dots \quad (100)$$

Proof:

The proof of this theorem is along the same lines as that of the previous theorem and so is omitted.

In many applications one usually substitute $\alpha_n \xi_n + \beta_n - \theta$ (ξ_n is the normalised order statistic of the case A, B or C) for $d_n \zeta_n$ and $e_n \zeta_n$ in (82) and (86) respectively while trying to obtain a V-statistic. In such cases it will be necessary to verify that $n \alpha_n^6 \rightarrow 0$ and $\frac{\beta_n - \theta}{\alpha_n} \rightarrow 0$ as $n \rightarrow \infty$ when (79) holds or that $n \alpha_n^2 \rightarrow \text{constant}$ and $\frac{\beta_n - \theta}{\alpha_n} \rightarrow 0$ as $n \rightarrow \infty$ when (83) and (84) hold. If the order statistic is of the case A and $F(x)$ possesses a non zero density function $f(x)$ at θ which is continuous in a neighbourhood of θ , then $\beta_n = \theta$ and $\alpha_n = \frac{1}{\sqrt{n} f(\theta)}$ so that the verifications just stated need not be carried out.

3.7. Alternative proofs and generalisations of the theorems of Heffding concerning mixtures of order statistics.

It is shown in this section that theorem 2.3.9 provides an alternative proof of the theorems 7.1, 7.4 of Hoeffding (1953). As stated earlier our method is elegant and permits of easy generalisations. We end this section with an application of our results.

The notations adopted here will conform to those of section 2. We have to define further quantities for the purposes of generalising some of our results.

Let $(x_1, y_1), \dots, (x_n, y_n)$ be n independent observations on a random variable (X, Y) with distribution function $G(x, y)$. The marginal distribution of X will be denoted by $F(x)$. We rearrange the sample with the x 's in the increasing order of magnitude as

$$(x_{(1)}^{(n)}, y_{(1)}^{(n)}), \quad \dots \quad (x_{(n)}^{(n)}, y_{(n)}^{(n)})$$

The notation adopted here being symmetric in x and y may be a bit confusing. We remark that we will never have to order the y 's so that $(x_{(r)}^{(n)}, y_{(r)}^{(n)})$ will always represent the observation whose x -component has the r th rank. Once again we omit the subscripts (n) for the sake of brevity. We shall denote the distribution function of $(x_{(r)}, y_{(r)})$ by $G(x, y; r, n)$. The distribution function of $(x_{(r)}, y_{(r)}, x_{(s)}, y_{(s)})$ will be denoted by $G(x, y, z, w; r, s, n)$.

3.7.1. Theorem.

$$\frac{1}{n} \sum_{r=1}^n F(x; r, n) = F(x) \quad \dots \quad (101)$$

Proof:

Let the rank of X_1 be R , when X_1, \dots, X_n are arranged in increasing order. R is a random variable taking the values $1, \dots, n$ with probability $\frac{1}{n}, \dots, \frac{1}{n}$.

$$\begin{aligned} F(x) &= P \{ X_1 \leq x \} \\ &= E_R [P \{ X_1 \leq x \mid R = r \}] \\ &= \frac{1}{n} \sum_{r=1}^n F(x; r, n). \end{aligned}$$

3.7.2. Theorem.

$$\frac{1}{n} \sum_{r=1}^n G(x, y; r, n) = G(x, y) \quad \dots \quad (102)$$

Proof:

The proof is similar to that of Theorem 7.1.

3.7.3. Theorem.

$$\frac{1}{n(n-1)} \sum_{r \neq s} F(x, y; r, s, n) = F(x) \cdot F(y) \dots \quad (103)$$

Proof:

The theorem is proved in the same way as Theorem 7.1. by considering (R, S) the ranks of X_1 and X_2 .

3.7.4. Theorem.

If $\{m_n\}$ is a sequence of integers such that $m_n/n \rightarrow \pi$, $0 < \pi < 1$, $F(\theta) = \pi$, and $F(x)$ continuous at θ , then

$$\frac{1}{m_n} \sum_{r=1}^{m_n} F(x; r, n) \rightarrow F_{\pi}(x) \quad \dots \quad (104)$$

at every continuity point of $F_{\pi}(x)$, where

$$F_{\pi}(x) = \begin{cases} F(x)/\pi & x \leq \theta \\ 1 & x > \theta \end{cases} \quad \dots \quad (105)$$

Proof:

Let R be a random variable taking the values $1, \dots, m_n$ with equal probabilities. Consider the variable $Y_n = X_{(R)}^{(n)}$.

$$P \{ Y_n \leq x \} = \frac{1}{m_n} \sum_{r=1}^{m_n} F(x; r, n) \quad \dots \quad (106)$$

Let $F_p^*(x)$ be the distribution function

$$F_p^*(x) = \begin{cases} F(x)/p & \text{if } x \leq F^{-1}(p) \\ 1 & \text{if } x > F^{-1}(p) \end{cases}$$

Let $F_p^*(x; r, m_n)$ be the distribution function of the r th order statistic from a sample of size m_n from this distribution. Let

$Z_n = F(X_{(m_n+1)})$. Then

$$P \left\{ Y_n \leq x \mid Z_n = p \right\} = \frac{1}{m_n} \sum_{r=1}^{m_n} F_p^*(x; r, m_n).$$

By Theorem 7.1 ^{therefore} we have

$$P \left\{ Y_n \leq x \mid Z_n = p \right\} = F_p^*(x).$$

Thus the conditional distributions of Y_n given $Z_n = p$ converges to $F_p^*(x)$ in the UC* sense relative to $p \in (\pi - \epsilon, \pi + \epsilon)$ for each sufficiently small ϵ . We also know that $Z_n \rightarrow p$ in probability. Thus from theorem 2.3.9 the joint distribution function of Y_n and Z_n converges in law, and hence the marginal distribution of Y_n converges in law to $F_\pi(x)$. Hence the theorem.

The following theorems are proved in the same way as theorem 7.4.

3.7.5. Theorem.

Let $\{m_n\}$ be a sequence of integers such that $\frac{m_n}{n} \rightarrow \pi$, $0 < \pi < 1$. $F(\theta) = \pi$ and $G(x, y)$ is continuous at θ for each y .

Then

$$\frac{1}{m_n} \sum_{r=1}^{m_n} G(x, y; r, n) \text{ converges in law to } G_\pi^*(x, y) \dots (107)$$

where

$$G_\pi^*(x, y) = \begin{cases} \frac{G(x, y)}{\pi} & \text{if } x \leq \theta \\ \frac{G(\theta, y)}{\pi} & \text{if } x > \theta \end{cases} \dots (108)$$

3.7.6. Theorem.

Let $\{k_n\}$, $\{l_n\}$ be two sequences of integers such that $1 \leq k_n \leq l_n \leq n$ and $\frac{k_n}{n} \rightarrow \pi_1$, $\frac{l_n}{n} \rightarrow \pi_2$, $0 < \pi_1, \pi_2 < 1$.

$F(\theta_1) = \pi_1$ and $F(\theta_2) = \pi_2$ and $F(x)$ continuous at θ_1 and θ_2 .

Then

$$\frac{1}{k_n n^{-k_n}} \sum_{r=1}^{k_n} \sum_{\substack{s=1 \\ s \neq r}}^n F(x, y; r, s, n) \text{ converges in law to } F_{\pi_1, \pi_2}^*(x, y) \quad \dots \quad (109)$$

where

$$F_{\pi_1, \pi_2}^*(x, y) = \begin{cases} \frac{F(x)F(y)}{\pi_1 \pi_2} & x, y \leq \theta_1 \\ \frac{F(y)}{\pi_2} & \theta_1 < x, y \leq \theta_2 \\ 1 & x, y > \theta_2 \end{cases} \quad \dots \quad (110)$$

The several theorems that we have given above serve to illustrate our techniques. We give below, in brief, an application of these theorems.

Suppose that the random variables described at the beginning of this section possesses finite second moments, i.e. $E(X^2) < \infty$ and $E(Y^2) < \infty$.

Let $\{m_n\}$ be a sequence of integers such that $\frac{m_n}{n} \rightarrow \pi$, $0 < \pi < 1$. Let $F(\theta) = \pi$ and $G(x, y)$ be continuous at θ for each y .

Let

$$E(X | X \leq \theta) = \mu, \quad E(Y | X \leq \theta) = \nu$$

$$V(X | X \leq \theta) = \sigma^2, \quad V(Y | X \leq \theta) = \tau^2 \quad \dots \quad (111)$$

$$\text{and cov}(X, Y | X \leq \theta) = \rho \sigma \tau$$

Let

$$\bar{x} = \frac{x_{(1)} + \dots + x_{(m_n)}}{m_n}$$

$$\bar{y} = \frac{y_{(1)} + \dots + y_{(m_n)}}{m_n}$$

$$s^2 = \frac{1}{m_n} \sum_{i=1}^{m_n} x_{(i)}^2 - \bar{x}^2 \quad \dots \quad (112)$$

$$t^2 = \frac{1}{m_n} \sum_{i=1}^{m_n} y_{(i)}^2 - \bar{y}^2$$

$$rst = \frac{1}{m_n} \sum_{i=1}^{m_n} x_{(i)} y_{(i)} - \bar{x} \bar{y}$$

Then these theorems can be used in the same way as in Hoeffding (1953)

to show that

$$\bar{x} \rightarrow \mu, \quad \bar{y} \rightarrow \nu, \quad s^2 \rightarrow \sigma^2, \quad t^2 \rightarrow \tau^2 \text{ and } rst \rightarrow \rho \sigma \tau \text{ in probability } \dots (113)$$

It has often been observed that there are some problems where fixed interval analysis becomes conceptually meaningless. For an important class of such situations, Professor P.C. Mahalanobis (1958a), (1958b), (1960) introduced a new method, called fractile analysis which also does away with the customary assumptions of regression analysis mentioned above. This method is described in section 2.

These two methods, fixed interval analysis and fractile analysis, throw up a large number of new and interesting problems in distribution theory, inference etc. The theoretical aspects of these problems are discussed in section 5, in this chapter. The limiting distributions required in this connection are derived in section 4 (some of these are found in Sethuraman (1961)) by making use of the theorems developed in Chapter II. For recent works in this direction we should refer to Takeuchi (1961) who has derived a theorem like ^{corollary} 4.11, but under more restrictive conditions. In the next chapter we shall discuss the practical aspects of these problems and provide tables necessary for their use.

Generalisations of the two methods, fixed interval analysis and fractile analysis are made in section 6. These throw open a large number of difficult and unsolved problems in stochastic processes. We end this chapter with the hope that these problems will be solved soon and thus throw more light on the methods of fixed interval

analysis and fractile analysis. We also indicate a method that will yield fruitful results in this connection. We then briefly summarise the results of Parthasarathy and Bhattacharya (1961) who have used this technique with success in connection with a related problem.

4.2. The methods of fixed interval analysis and fractile analysis.

One of the important problems of statistics is the study of the relationship of one variable Y with another variable X and of the comparison of such relationships ^{among} ~~for~~ ^{populations} ~~problems~~. These are usually made through the study of the regression function of Y on X , i.e. through the study of the function $E(Y | X = x)$. A common problem met with in practice is the comparison of the regression functions in two populations (Y, X) and (Y', X') . We can give illustrations for this problem from practically every applied field of Statistics. We will however content ourselves with one example which we will refer to, for purposes of illustration, in the sequel. Data on a pair of random variables (Y, X) and (Y', X') are available. Y and X refer to the consumption of milk (volume) and total expenditure per month, respectively, of an individual in a population P . Y' and X' refer to the same variates in a different population P' . The problem is to compare the patterns of relationship of the consumption of milk with the total expenditure in the two populations.

The problem just stated can be expressed in symbols as follows:

We wish to test whether or not

$$\lambda(x) = \lambda'(x) \text{ for all } x \quad \dots \quad (1)$$

where

$$\lambda(x) = E(Y \mid X = x) \quad \dots \quad (2)$$

$$\lambda'(x) = E(Y' \mid X' = x) \quad \dots \quad (3)$$

This is the sort of regression problem we shall be concerned with in this chapter. The classical method of approaching this problem consists in assuming that the regression functions $\lambda(x)$ and $\lambda'(x)$ are of a certain algebraic form, completely determined except for a finite number of 'parameters', say, a polynomial, a trigonometric series or the like. The problem of testing the equality of the regression functions then reduces to the problem of testing the equality of these parameters. The difficulty with these methods is that the assumption of algebraic model for the two regression functions is too great an over simplification of the actual situations. Further, it is very difficult to test whether any particular model fits in a given practical set up. We will describe in this section two methods, fixed interval analysis and fractile analysis, which do not involve the assumptions of the usual regression models. In both these methods we compare the two regression functions intervalwise instead of pointwise, where the intervals are largely at our choice.

In fixed interval analysis we fix to start with g fixed intervals

$$(a_0, a_1], (a_1, a_2], \dots, (a_{g-1}, a_g] \dots \quad (4)$$

by means of the $(g+1)$ constants

$$a_0, a_1, \dots, a_g \dots \quad (5)$$

satisfying

$$-\infty = a_0 < a_1 < \dots < a_{g-1} < a_g = +\infty \dots \quad (6)$$

We then define the regression $\bar{y}(a, b)$ of Y on X in an interval $(a, b]$ by the relation

$$\bar{y}(a, b) = E(Y \mid a < X \leq b) \dots \quad (7)$$

Let

$$\bar{y}(a_{i-1}, a_i) = \bar{y}_i, \quad i = 1, \dots, g \dots \quad (8)$$

The quantities \bar{y}'_i are defined in the same way for the random variable (Y', X') . In fixed interval analysis we test the hypothesis that

$$\bar{y} = \bar{y}' \dots \quad (9)$$

where

$$\bar{y} = (\bar{y}_1, \dots, \bar{y}_g), \quad \bar{y}' = (\bar{y}'_1, \dots, \bar{y}'_g) \dots \quad (10)$$

It is obvious that such a hypothesis becomes meaningful only when the intervals $(a_0, a_1], \dots, (a_{g-1}, a_g]$ of (4) are chosen carefully. For instance, in our illustrative example, let the populations P and P' correspond to two constituent states of the Indian Union at the same point in time. Several comparable social strata can be defined in the two populations by a suitable choice of the intervals of (4). Such strata might correspond, e.g., to incomes classified as 'lower class', 'lower middle class', etc. In this case the hypothesis (9) is meaningful and fixed interval analysis is appropriate to the situation.

It should be noted that the use of a large number of strata would mean that the hypothesis $\bar{y} = \bar{y}'$ is practically equivalent to the hypothesis $\lambda(x) = \lambda'(x)$.

The practical details of the method are as follows. We estimate the quantities \bar{y} and \bar{y}' suitably and test for their difference. Let

$$(y_1, x_1), \dots, (y_n, x_n) \dots \quad (11)$$

and

$$(y'_1, x'_1), \dots, (y'_{n'}, x'_{n'}) \dots \quad (12)$$

be two independent samples S and S' from the populations P and P' respectively. Let n_i be the number of observations of the sample S with x -components in the interval $(a_{i-1}, a_i]$, $i = 1, \dots, g$.

Let

$$\bar{v}_i = \frac{\sum_{a_{i-1} < x_r \leq a_i} y_r}{n_i}, \quad i = 1, \dots, g \quad \dots \quad (13)$$

where the summation is made over all y_r 's for which x_r 's lie in $(a_{i-1}, a_i]$. We calculate \bar{v}_i , $i = 1, \dots, g$ in a similar way from the sample S' . The vectors \bar{v} and \bar{v}' are estimates of \bar{y} and \bar{y}' . We define several measures of divergence between the sample estimates for the two regressions.

$$\bar{D} = \sum_{i=1}^g |\bar{v}_i - \bar{v}'_i| \quad \dots \quad (14)$$

$$\bar{\Delta} = \sum_{i=1}^g (\bar{v}_i - \bar{v}'_i)^2 \quad \dots \quad (15)$$

$$\bar{\Gamma} = (\bar{v} - \bar{v}') B (\bar{v} - \bar{v}')' \quad \dots \quad (16)$$

where B is some positive definite matrix. Large values of these statistics will form the corresponding rejection region for testing the hypothesis in (9). It is an important problem to determine the distributions of these statistics, atleast in large samples, and set up significance points for these statistics.

We now proceed to remark that there are many problems where fixed interval analysis becomes conceptually meaningless or partially so. As an example let us suppose that in our practical example, the

populations P and P' correspond to two different states having different currencies. After setting up interval limits to the total expenditure reflecting different social groups in one population, we may find it very difficult to demarcate comparable limits in the currency of the second population. The official exchange rate cannot be used here since it does not reflect the actual purchasing power of the two currencies. The real exchange rate that does this is not easily available. We have thus presented a typical case of the general situation where X and X' are not comparable and where fixed interval analysis is not useful.

In the above example we can safely assume that the total expenditure is a monotonic function of the socio-economic level of an individual. Thus, though the total expenditures in the two countries are not directly comparable they are monotonically related. Groupings based on the ranks of X and X' will be comparable in a meaningful way. Professor P. C. Mahalanobis made use of this fact in proposing a new method called fractile analysis for such situations.

We describe this method after defining some quantities. Let $\theta_1, \theta_2, \dots, \theta_{g-1}$ be the $\frac{1}{g}$ th, $\frac{2}{g}$ th, \dots , $\frac{g-1}{g}$ th quantiles (or fractiles) of the distribution of X . Let $\theta_0 = -\infty$, $\theta_g = +\infty$. Let $\theta'_0, \theta'_1, \dots, \theta'_{g-1}, \theta'_g$ be the corresponding quantities for the distribution of X' . Let

$$\nu_i = \bar{\nu}(\theta_{i-1}, \theta_i) \quad i = 1, \dots, g \quad \dots \quad (17)$$

$$\nu_i' = \bar{\nu}'(\theta_{i-1}', \theta_i') \quad i = 1, \dots, g \quad \dots \quad (18)$$

We see that the intervals $(\theta_0, \theta_1), (\theta_1, \theta_2) \dots (\theta_{g-1}, \theta_g)$ of X represent the lowest $100/g$ percent section, the second lowest $100/g$ percent section the highest $100/g$ percent section, respectively, of population P . A similar interpretation can be given for the intervals $(\theta_0', \theta_1'), (\theta_1', \theta_2'), \dots (\theta_{g-1}', \theta_g')$ of X' in the second population P' . Thus these intervals are comparable in a very important sense, although the X values are different. The method of fractile analysis consists in testing the hypothesis

$$\mathcal{Z} = \mathcal{Z}' \quad \dots \quad (19)$$

where

$$\mathcal{Z} = (\nu_1, \dots, \nu_g), \quad \mathcal{Z}' = (\nu_1', \dots, \nu_g') \quad \dots \quad (20)$$

The practical method adopted is as follows. Let S and S' be the samples from the populations P and P' respectively, as in (11), (12). Rearrange the observations in the sample S so that the x 's are in the increasing order of magnitude thus

$$(y_{(1)}, x_{(1)}), (y_{(2)}, x_{(2)}), \dots (y_{(n)}, x_{(n)}) \dots \quad (21)$$

with $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)} \dots$ (22)

We repeat the remark made in Section 3.7. We will stick to this notation despite its symmetry in y and x and there will be no confusion for we will never have occasion to order the y 's. This notation will be carried till the end of this work. Let $n = m \cdot g$ where m and g are integers. We define the quantities

$$v_i = \frac{\sum_{(i-1)m < r \leq im} y_{(r)}}{m}, \quad i = 1, \dots, g \quad \dots (23)$$

Quantities v'_i , $i = 1, \dots, g$ are defined in a similar way from S' . Vectors \underline{v} and \underline{v}' are estimates of \underline{y} and \underline{y}' . Suitable measures of divergence are then defined. Professor P.C. Mahalanobis defined one such measure called separation, as follows. Plot the ordinates v_1, v_2, \dots, v_g corresponding to the equidistant points 1, 2, \dots , g . Join the successive points by straight lines; the curve G thus obtained is called the fractile graph of S . The fractile graph G' of S' is obtained in a similar way and drawn on the same paper and to the same scale. The area A between these two graphs, and the ordinates at 1 and g called the separation, is a measure of divergence between the two sample regressions. The algebraic expression for A is

$$A = \sum_{i=1}^{g-1} \left[\left\{ \frac{1}{2} |v_i - v'_i| + \frac{1}{2} |v_{i+1} - v'_{i+1}| \right\} - \delta(v_i - v'_i, v_{i+1} - v'_{i+1}) \frac{|(v_i - v'_i)(v_{i+1} - v'_{i+1})|}{|v_i - v'_i| + |v_{i+1} - v'_{i+1}|} \right] \quad (24)$$

$$\text{where } \delta(a, b) = \begin{cases} 0 & \text{if } ab > 0 \\ 1 & \text{if } ab < 0 \end{cases} \quad \dots \quad (25)$$

Some other measures of divergence which have been considered are

$$D = \sum_{i=1}^g |v_i - v'_i|, \quad \dots \quad (26)$$

$$\Delta = \sum_{i=1}^g (v_i - v'_i)^2, \quad \dots \quad (27)$$

$$\Gamma = (\underline{v} - \underline{v}') B (\underline{v} - \underline{v}')', \quad \dots \quad (28)$$

where B is some positive definite matrix. Large values of these statistics form the corresponding critical regions for testing the hypothesis in (19). We repeat the remark made earlier, that at least the large sample distributions of these statistics should be determined before using these measures for testing purposes. Till now only descriptive methods are available for the methods of fixed interval analysis and fractile analysis; for examples see Mahalanobis⁽¹⁹⁶⁰⁾,

Das (1960), Das and Sharma (1960).

We find the requisite limiting distributions in section 4 and these can now be used in practice.

Let us add a few points bringing out the special character of the method of fractile analysis in comparison with the method of fixed interval analysis. For applying the method of fixed interval analysis the random variables X and X' must be directly comparable and some intervals resembling strata should be formed. We then compare the regression in these intervals. The method of fractile analysis is applicable to the more difficult situation where X and X' are not directly comparable but are monotonically related to a common character that X and X' are supposed to measure. This, then, is the general set up where fractile analysis can be used. Such situations have been encountered in Econometrics, Psychometry, Demography etc., and fractile analysis has been applied, though, as yet only in a descriptive way. See Mahalanobis (1960), Das (1960), Das and Sharma (1960), Som (1960). Fractile analysis is now being utilised in the National Sample Survey of India on a large scale.

Several modifications of the method of fractile analysis can be made. After ranking the individuals we can take groups with varying proportions instead of equal proportions. Thus a group already formed may be split up into several sub-groups for a more detailed analysis of the regression in that group. This has been employed in

Mahalanobis (1960). For calculating the sample estimate of χ , the median, mode or some other suitable characteristic can be taken instead of the mean. These may be easier to compute in practice. This has been stated in Mahalanobis (1958a). We can thus go on multiplying the number of modifications we could make.

In practice, it happens that frequently we take several and independent sub-samples S_1, \dots, S_k from a population P instead of just one sample. These are called interpenetrating sub-samples and their usefulness in a large number of situations has been recognised. See Mahalanobis (1946), Yates (1949), Gosh (1949), Lahiri (1954), (1957). If interpenetrating sub-samples are taken from both the populations P and P' we can get measures, D , D_1 and D_2 , of divergence, from the samples, measuring the divergence between the combined samples of P and P' , within the sub-samples of P and within the sub-samples of P' , respectively. Thus D would be the between divergence whereas D_1 and D_2 would be the within divergences. These terminologies are analogous to those in analysis of variance problems and are self-explanatory. Suggestive as they are they permit us to develop some tests of at least a descriptive nature. See Mahalanobis (1958a), (1960).

Finally we add that the quantities calculated in the methods of fixed interval analysis and fractile analysis can be suitably modified to estimate relative concentration curves and relative

concentration ratios. The utility of these is well known. See for example, Davis (1941), Roy, Chakravarti and Laha (1959), Iyengar (1960).

We now give a brief description of the method of obtaining the concentration curve etc. The following symbols will be used:

$$\bar{v}_i = n_1 \bar{v}_1 + n_2 \bar{v}_2 + \dots + n_i \bar{v}_i; \quad i = 1, \dots, g \quad (29)$$

$$\bar{v}_0 = 0, \quad \bar{v}_g = \bar{v} \quad \dots \quad \dots \quad \dots \quad (30)$$

$$\bar{q}_i = \bar{v}_i / \bar{v}; \quad i = 0, 1, \dots, g \quad \dots \quad (31)$$

$$\bar{p}_i = (n_1 + n_2 + \dots + n_i) / n; \quad i = 1, \dots, g \quad (32)$$

$$\bar{p}_0 = 0 \quad \dots \quad \dots \quad \dots \quad (33)$$

$$v_i = (v_1 + \dots + v_i); \quad i = 1, \dots, g \quad \dots \quad (34)$$

$$v_0 = 0, \quad v_g = v \quad \dots \quad \dots \quad \dots \quad (35)$$

$$q_i = v_i / v; \quad i = 0, 1, \dots, g \quad \dots \quad \dots \quad (36)$$

$$p_i = i / g; \quad i = 0, \dots, g \quad \dots \quad \dots \quad (37)$$

When the method of fixed interval analysis is used an estimate of the relative concentration curve of Y on X is obtained by plotting the points

$$(\bar{p}_0, \bar{q}_0), (\bar{p}_1, \bar{q}_1), \dots, (\bar{p}_g, \bar{q}_g) \quad \dots \quad (38)$$

and joining successive points by straight lines. The relative concentration ratio is estimated by

$$\bar{C} = \frac{1}{\bar{V}} \left[\sum n_1 \bar{v}_1 (\bar{P}_{1-1} + \bar{P}_1) \right] - 1 \quad \dots \quad (39)$$

When the method of fractile analysis is used an estimate of the relative concentration curve of Y on X is obtained by plotting the points

$$(p_0, q_0), (p_1, q_1), \dots (p_g, q_g) \quad \dots \quad (40)$$

and joining successive points by straight lines. The relative concentration ratio is estimated by

$$C = \frac{2}{gV} \sum i v_i - \frac{g+1}{g} \quad \dots \quad (41)$$

4.3. Notation, definitions etc.

In this section we develop the notations etc., to be used in sections 4 and 5. Since we are developing the theories of fixed interval analysis and fractile analysis side by side and since they are similar to one another in many respects, the notations employed will also be similar. Whenever possible we will distinguish the quantities involved in fixed interval analysis by a bar $\bar{\quad}$.

Let (Y, X) be a random variable on the Euclidean plane with distribution function $G(y, x)$. The distribution function of X

will be denoted by $F(x)$.

$$\text{Let } (y_1, x_1), (y_2, x_2), \dots (y_n, x_n) \dots \quad (42)$$

be a sample S of n independent observations on (Y, X) . Let us arrange the observations according to the increasing order of magnitude of the x'_i , thus:

$$(y_{(1)}, x_{(1)}), (y_{(2)}, x_{(2)}), \dots (y_{(n)}, x_{(n)}) \dots \quad (43)$$

The fixed interval method of analysis involves stratifying the population into g strata. Let these strata be formed by the predetermined constants $a_0, a_1, \dots a_g$ satisfying

$$-\infty = a_0 < a_1 < \dots < a_{g-1} < a_g = +\infty \quad \dots \quad (44)$$

These constants introduce g strata in the domain of (Y, X) as follows: The r th stratum ($r = 1, \dots g$) contains all (y, x) with $a_{r-1} < x \leq a_r$. We define

$$\pi_r = F(a_r) - F(a_{r-1}); \quad i = 1, \dots, g \quad \dots \quad (45)$$

throughout the discussion on fixed interval analysis we shall assume that

$$\pi_r > 0, \quad r = 1, \dots, g \quad \dots \quad (46)$$

Let us define the means, variances and covariances of Y and X (these are assumed to exist throughout this discussion) in these g strata

$$\left. \begin{aligned}
 \bar{\mu}_r &= E(X \mid a_{r-1} < X \leq a_r); & \bar{\nu}_r &= E(Y \mid a_{r-1} < X \leq a_r) \\
 \bar{\sigma}_r^2 &= V(X \mid a_{r-1} < X \leq a_r); & \bar{c}_r^2 &= V(Y \mid a_{r-1} < X \leq a_r) \\
 \bar{\sigma}_r \bar{\sigma}_r \bar{c}_r &= \text{cov}^*(Y, X \mid a_{r-1} < X \leq a_r); & r &= 1, \dots, g.
 \end{aligned} \right\} \dots (47)$$

Let n_1, n_2, \dots, n_g be the number of observations in the sample S in the 1st, 2nd, ..., gth stratum, respectively, introduced by the constants in (44). We shall denote the proportion of observations in these strata by p_1, p_2, \dots, p_g , i.e.

$$p_i = n_i / n; \quad i = 1, \dots, g \quad \dots \quad (48)$$

The means, variances and covariances the sample S in these strata are defined as follows:

* $\text{cov}(\xi, \eta)$ represents the covariance between ξ and η .

$$\left. \begin{aligned}
 \bar{u}_i &= \sum_{a_{i-1} < x_r \leq a_i} x_r / n_i; \quad \bar{v}_i = \sum_{a_{i-1} < x_r \leq a_i} y_r / n_i \\
 \bar{s}_i^2 &= \sum_{a_{i-1} < x_r \leq a_i} (x_r - u_i)^2 / n_i; \quad \bar{t}_i^2 = \sum_{a_{i-1} < x_r \leq a_i} (y_r - v_i)^2 / n_i \\
 \bar{r}_i \bar{s}_i \bar{t}_i &= \sum_{a_{i-1} < x_r \leq a_i} (x_r - u_i)(y_r - v_i) / n_i; \quad i = 1, \dots, g.
 \end{aligned} \right\} (49)$$

In the theorems of the next section will be interested in the limiting distribution of the following statistics.

$$\left. \begin{aligned}
 \bar{\xi}_i(n) &= \sqrt{n} (\bar{u}_i - \mu_i); \quad \bar{\eta}_i(n) = \sqrt{n} (\bar{v}_i - \nu_i); \\
 \bar{\zeta}_i(n) &= \sqrt{n} (p_i - \pi_i); \quad i = 1, \dots, g
 \end{aligned} \right\} \dots \quad (50)$$

$$\begin{aligned}
 \bar{\Sigma}(n) &= (\bar{\xi}_1(n), \dots, \bar{\xi}_g(n)); \quad \bar{\Upsilon}(n) = (\bar{\eta}_1(n), \dots, \bar{\eta}_g(n)); \\
 \bar{\Xi}(n) &= (\bar{\zeta}_1(n), \dots, \bar{\zeta}_g(n)). \quad \dots \quad \dots \quad (51)
 \end{aligned}$$

In problems connected with fractile analysis we shall assume the following:

The distribution function $G(y, x)$ admits of a density function $g(y, x)$ which is continuous and which does not vanish. } (52)

The density function of $F(x)$ will be denoted by $f(x)$.

In the fractile method of analysis the population is to be divided into a previously assigned number, g , of strata thus.

Let θ_i be defined by

$$F(\theta_i) = i/g \quad i = 1, \dots, g-1 \quad \dots \quad (53)$$

$$\theta_0 = -\infty, \quad \theta_g = +\infty \quad \dots \quad \dots \quad (54)$$

Then the r^{th} stratum is defined as the region of all points (y, x) with

$$\theta_{r-1} < x \leq \theta_r, \quad r = 1, \dots, g.$$

The means, variances and covariances (these are assumed to exist throughout this discussion) in these strata are defined by

$$\left. \begin{aligned} \mu_r &= E(X \mid \theta_{r-1} < X \leq \theta_r); \quad \nu_r = E(Y \mid \theta_{r-1} < X \leq \theta_r); \\ \sigma_r^2 &= V(X \mid \theta_{r-1} < X \leq \theta_r); \quad \tau_r^2 = V(Y \mid \theta_{r-1} < X \leq \theta_r); \\ \rho_r \sigma_r \tau_r &= \text{cov}(Y, X \mid \theta_{r-1} < X \leq \theta_r); \quad r = 1, \dots, g. \end{aligned} \right\} (55)$$

We also require the regression function

$$E(Y \mid X = x) = \lambda(x) \quad \dots \quad (56)$$

$$\lambda(\theta_i) = \lambda_i, \quad i = 1, \dots, g-1 \quad \dots \quad (57)$$

$$\lambda_0 = 0, \quad \lambda_g = 0 \quad \dots \quad \dots \quad \dots \quad (58)$$

The corresponding quantities are defined from the sample S.

It is assumed that $n = m \cdot g$ where m is an integer.

$$\begin{aligned}
 u_i &= \frac{\sum_{(i-1)m < r \leq im} x(r)}{m}; \\
 v_i &= \frac{\sum_{(i-1)m < r \leq im} y(r)}{m}; \\
 s_i^2 &= \frac{\sum_{(i-1)m < r \leq im} (x(r) - u_i)^2}{m}; \quad t_i^2 = \frac{\sum_{(i-1)m < r \leq im} (y(r) - v_i)^2}{m}; \\
 r_i s_i t_i &= \frac{\sum_{(i-1)m < r \leq im} (x(r) - u_i)(y(r) - v_i)}{m}; \quad i=1, \dots, g.
 \end{aligned} \tag{59}$$

In the next section we will be interested in the limiting distribution of the following statistics.

$$\left. \begin{aligned}
 \xi_i(n) &= \sqrt{m} (u_i - \mu_i); \quad i = 1, \dots, g \\
 \eta_i(n) &= \sqrt{m} (v_i - \nu_i); \quad i = 1, \dots, g
 \end{aligned} \right\} \dots \tag{60}$$

$$\zeta_i(n) = \sqrt{n} (x_{(im+1)} - \theta_i); \quad i = 1, \dots, g-1 \quad \dots \tag{61}$$

$$\begin{aligned}
 \xi(n) &= (\xi_1(n), \dots, \xi_g(n)); \quad \eta(n) = (\eta_1(n), \dots, \eta_g(n)); \\
 \zeta(n) &= (\zeta_1(n), \dots, \zeta_{g-1}(n)). \quad \dots \quad \dots \tag{62}
 \end{aligned}$$

In the following sections we shall have occasion to consider another random variable (Y', X') constituting a population P' . The constants for this population will be obtained by adding a mark to those of P . In the same way the quantities for any sample with some suffix will be obtained by adding that suffix to the quantities of the sample S .

4.4. Limit distributions.

In this section we state and prove some results concerning the limit distributions of the statistics entering in (51) and (62). Some of these results are found in Sethuraman (1961).

4.4.1. Theorem. Let condition (45) hold. As $n \rightarrow \infty$, the sequence of random variables $(\bar{\xi}(n), \bar{\eta}(n), \bar{\zeta}(n))$ converges weakly to a random variable $(\bar{\xi}, \bar{\eta}, \bar{\zeta})$ with a multivariate normal distribution.

Proof:

We shall first prove the following lemma.

4.4.2. Lemma. The conditional distribution function of $(\bar{\xi}(n), \bar{\eta}(n))$ given $\bar{\zeta}(n) = \underline{z}$ converges in law to a multivariate normal distribution in the UC* sense relative to \underline{z} in any bounded interval I . The limiting conditional means, variances and covariances are independent of \underline{z} .

Proof:

The proof of this lemma proceeds on the same lines as lemma 3.4.2 except that we have to use a generalised version of theorem 3.3.3.

It is easy to see that as soon as $\bar{\xi}(n)$ is fixed at \underline{g} , we can assume that there are $[n \kappa_r + \sqrt{n} z_r]$ observations in the r^{th} stratum defined by (44), $r = 1, \dots, g$ and that these observations are independent. Thus $(\bar{\xi}(n), \bar{\eta}(n))$ are the normalised means from independent samples from g bivariate distributions. Now we observe that $[n \kappa_r + \sqrt{n} z_r] \rightarrow \infty$ uniformly in $\underline{g} \in I$ for $r = 1, \dots, g$ and that $\frac{n \sigma_r^2}{[n \kappa_r + \sqrt{n} z_r]}$ etc. tends to $\frac{\sigma_r^2}{\kappa_r}$ etc. uniformly in $\underline{g} \in I$ and for $r = 1, \dots, g$. The lemma now follows from the easily obtainable generalisation of theorem 3.3.3 for multivariate random variables.

The theorem 4.1 is immediate from the lemma 4.2 the well known fact that $\bar{\xi}(n)$ has a limiting normal distribution and the theorem 2.3.9.

We now state several corollaries of the above theorem.

4.4.3 Corollary. The distribution of $(\bar{\xi}, \bar{\eta}, \bar{\zeta})$ is multivariate normal with mean vector 0 and variance covariance matrix

$$\begin{pmatrix} \bar{\Sigma} \Pi' & \bar{E} \Pi' & 0 \\ \bar{E} \Pi' & \bar{T} \Pi' & 0 \\ 0 & 0 & \bar{K} \end{pmatrix} \begin{matrix} (g) \\ (g) \\ (g) \end{matrix} \dots (63)$$

$$\text{where } \bar{\Sigma} = \text{diag}(\bar{\sigma}_1^2, \dots, \bar{\sigma}_g^2) \dots \dots \dots \quad (64)$$

$$\bar{\Gamma} = \text{diag.} (\bar{\tau}_1^2, \dots, \bar{\tau}_g^2) \dots \dots \quad (65)$$

$$\bar{\Xi} = \text{diag.} (\bar{\sigma}_1 \bar{\sigma}_1 \bar{\tau}_1, \dots, \bar{\sigma}_g \bar{\sigma}_g \bar{\tau}_g) \dots \quad (66)$$

$$\bar{\Lambda} = \text{diag} (\pi_1, \dots, \pi_g) \dots \dots \quad (67)$$

$$\bar{K} = \begin{pmatrix} \pi_1(1-\pi_1) & \pi_2(1-\pi_2) & \dots & \pi_1(1-\pi_g) \\ \pi_1(1-\pi_2) & \pi_2(1-\pi_2) & \dots & \pi_2(1-\pi_g) \\ \dots & \dots & \dots & \dots \\ \pi_1(1-\pi_g) & \pi_2(1-\pi_g) & \dots & \pi_g(1-\pi_g) \end{pmatrix} \dots \quad (68)$$

4.4.4. Corollary: The distribution of $(\bar{\Sigma}, \bar{\Gamma})$ is multivariate normal with mean vector 0 and variance covariance matrix

$$\begin{pmatrix} \bar{\Sigma} \bar{\Lambda}^{-1} & | & \bar{\Xi} \bar{\Lambda}^{-1} \\ \hline \bar{\Xi} \bar{\Lambda}^{-1} & | & \bar{\Gamma} \bar{\Lambda}^{-1} \end{pmatrix} \begin{matrix} (g) \\ (g) \end{matrix}$$

4.4.5. Corollary: The distribution of $\bar{\eta}$ is multivariate normal with mean vector 0 and variance covariance matrix where

$$\bar{\Lambda} = \bar{\Gamma} \bar{\Xi}^{-1} \dots \dots \quad (69)$$

* $\text{diag} (b_1, \dots, b_k)$ stands for the diagonal matrix $\begin{pmatrix} b_1 & 0 & \dots & 0 \\ 0 & b_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_k \end{pmatrix}$.

4.4.6. Corollary: Let

$$\bar{\xi}_i^0(n) = \sqrt{p_1} \bar{\xi}_i(n), \quad \bar{\eta}_i^0(n) = \sqrt{p_1} \bar{\eta}_i(n), \quad i=1, \dots, g \quad (70)$$

Then $(\bar{\xi}_i^0(n), \bar{\eta}_i^0(n))$ tends weakly to a random variable $(\bar{\xi}_i^0, \bar{\eta}_i^0)$ as $n \rightarrow \infty$ and the limiting distribution is multivariate normal with mean vector 0 and variance covariance matrix

$$\begin{pmatrix} \bar{\Sigma} & \bar{K} \\ \bar{K} & \bar{\Gamma} \end{pmatrix} \begin{matrix} (g) \\ (g) \end{matrix}$$

4.4.7. Theorem. Let condition (52) hold. As $n \rightarrow \infty$

(i.e. $n \rightarrow \infty$) the sequence of random variables $(\xi(n), \eta(n), \zeta(n))$ converges weakly to a random variable (ξ, η, ζ) with a multivariate normal distribution.

Proof:

We will prove the following lemma first.

4.4.8. Lemma. The conditional distribution function of

$(\xi(n), \eta(n))$ given $\zeta(n) = z$ converges in law to the distribution function of a multivariate normal with a mean vector linear in z and a variance covariance matrix independent of z .

Proof:

For shortening the proof and simplifying the algebra we present the proof for the case $g = 2$ only. The proof for the general case

proceeds on the same lines.

Since $g = 2$, $n = 2m$. $\xi(n) = \zeta_1(n) = \sqrt{n}(x_{(m+1)} - \theta_1)$.
 When $\zeta_1(n)$ is fixed at z_1 , $x_{(m+1)}$ is fixed at $\theta_1 + \frac{z_1}{\sqrt{n}}$ and
 thus the conditional distribution of the sample is that of a sample
 of m independent observations on (Y, X) truncated to the region
 $(-\infty < x < \theta_1 + \frac{z_1}{\sqrt{n}}, -\infty < y < \infty)$, an independent sample of
 $n - m - 1 (= m - 1)$ independent observations on (Y, X) truncated to
 the region $(\theta_1 + \frac{z_1}{\sqrt{n}} < x < \infty, -\infty < y < \infty)$ and an independent
 observation on (Y, X) truncated to the region $(x = \theta_1 + \frac{z_1}{\sqrt{n}},$
 $-\infty < y < \infty)$. ($\xi_1(n)$, $\xi_2(n)$, $\eta_1(n)$, $\eta_2(n)$) are normalised means
 from these samples. Now it is easy to verify that

$$\begin{aligned} \sqrt{m} (\mu_1 - E(X | X < \theta_1 + \frac{z_1}{\sqrt{n}})) &\rightarrow \sqrt{2} z_1 f(\theta_1)(\theta_1 - \mu_1); \\ \sqrt{m} (\nu_1 - E(Y | X < \theta_1 + \frac{z_1}{\sqrt{n}})) &\rightarrow \sqrt{2} z_1 f(\theta_1)(\lambda_1 - \nu_1); \\ \sqrt{m} (\mu_2 - E(X | X > \theta_1 + \frac{z_1}{\sqrt{n}})) &\rightarrow -\sqrt{2} z_1 f(\theta_1)(\theta_1 - \mu_2); \\ \sqrt{m} (\nu_2 - E(Y | X > \theta_1 + \frac{z_1}{\sqrt{n}})) &\rightarrow -\sqrt{2} z_1 f(\theta_1)(\lambda_1 - \nu_2); \\ V(X | X < \theta_1 + \frac{z_1}{\sqrt{n}}) &\rightarrow \sigma_1^2; \quad V(Y | X < \theta_1 + \frac{z_1}{\sqrt{n}}) \rightarrow \tau_1^2; \\ V(X | X > \theta_1 + \frac{z_1}{\sqrt{n}}) &\rightarrow \sigma_2^2; \quad V(Y | X > \theta_1 + \frac{z_1}{\sqrt{n}}) \rightarrow \tau_2^2; \\ \text{cov}(Y, X | X < \theta_1 + \frac{z_1}{\sqrt{n}}) &\rightarrow \rho_1 \sigma_1 \tau_1; \quad \text{cov}(Y, X | X > \theta_1 + \frac{z_1}{\sqrt{n}}) \rightarrow \rho_2 \sigma_2 \tau_2; \\ (\frac{x_{(m+1)} - \mu_2}{\sqrt{m}}, \frac{y_{(m+1)} - \nu_2}{\sqrt{m}}) &\rightarrow 0; \text{ in probability} \end{aligned} \quad (71)$$

for each z_1 .

Using the generalised version of theorem 3.3.1 to include bivariate random variables and the above relations we see that the sequence of random variables $(\xi(n), \eta(n))$ given $\xi(n) = \bar{x}$ converges to a normal distribution with mean vector

$$\begin{aligned} &(\sqrt{2} z_1 f(\theta_1)(\theta_1 - \mu_1), -\sqrt{2} z_1 f(\theta_1)(\theta_1 - \mu_2), \sqrt{2} z_1 f(\theta_1)(\lambda_1 - \gamma_1), \\ &-\sqrt{2} z_1 f(\theta_1)(\lambda_1 - \gamma_2)) \quad \dots \quad \dots \quad (72) \end{aligned}$$

and variance covariance matrix

$$\begin{pmatrix} \sigma_1^2 & 0 & \rho_1 \sigma_1 \tau_1 & 0 \\ 0 & \sigma_2^2 & 0 & \rho_2 \sigma_2 \tau_2 & \dots \\ \rho_1 \sigma_1 \tau_1 & 0 & \tau_1^2 & 0 \\ 0 & \rho_2 \sigma_2 \tau_2 & 0 & \tau_2^2 \end{pmatrix} \quad (73)$$

Hence the lemma.

We know from Cramer (1946) and the theorem of Scheffé (1947), that under the condition (52), $\xi_m(n)$ converges strongly to multivariate normal distribution. The theorem 4.7 follows from this fact, the lemma 4.8 and the theorem 2.3.3.

We now state several corollaries.

4.4.9. Corollary: The sequence of random variables $(\xi^{(n)}, \eta^{(n)})$ given $\xi^{(n)} = \underline{x}$ converge weakly to a random variable with a multivariate normal distribution with mean vector

$$\underline{x} \begin{pmatrix} (A_1 & | & A_2) \\ \hline \text{---} & & \text{---} \\ \text{---} & & \text{---} \end{pmatrix}^{(g-1)} \dots \quad (74)$$

and variance covariance matrix

$$\begin{pmatrix} \Sigma & | & E \\ \hline \text{---} & & \text{---} \\ \text{---} & & \text{---} \end{pmatrix}^{(g)} \dots \dots \quad (75)$$

where $\Sigma = \text{diag. } (\sigma_1^2, \dots, \sigma_g^2) \dots \dots \quad (76)$

$\Gamma = \text{diag. } (\tau_1^2, \dots, \tau_g^2) \dots \dots \quad (77)$

$E = \text{diag. } (\rho_1 \sigma_1 \tau_1, \dots, \rho_g \sigma_g \tau_g) \dots \dots \quad (78)$

where

$$\left. \begin{aligned} (A_1)_{ii} &= \sqrt{g} f(\theta_i)(\theta_i - \mu_i), \quad i = 1, \dots, g-1, \\ (A_1)_{ii+1} &= -\sqrt{g} f(\theta_i)(\theta_i - \mu_{i+1}), \quad i = 1, \dots, g-1, \\ (A_1)_{ij} &= 0 \quad \text{for other terms} \end{aligned} \right\} \dots \quad (79)$$

$$\left. \begin{aligned} (A_2)_{ii} &= \sqrt{g} f(\theta_i)(\lambda_i - \nu_i), \quad i = 1, \dots, g-1, \\ (A_2)_{ii+1} &= -\sqrt{g} f(\theta_i)(\lambda_i - \nu_{i+1}), \quad i = 1, \dots, g-1, \\ (A_2)_{ij} &= 0 \quad \text{for all other terms} \quad \dots \end{aligned} \right\} \quad (80)$$

* $(B)_{ij}$ will denote the (i,j) th element of the matrix B.

4.4.10. Corollary: The distribution of (ξ, η, ζ) is multivariate normal with mean vector 0 and variance covariance matrix

$$\begin{pmatrix} \Sigma + A_1' K A_1 & E + A_1' K A_2 & A_1' K & (g) \\ \hline E + A_2' K A_1 & T + A_2' K A_2 & A_2' K & (g) \\ \hline K' A_1 & K' A_2 & K & (g-1) \end{pmatrix} \dots \quad (81)$$

where

$$(K)_{ij} = (K)_{ji} = \frac{i(g-j)}{g} \cdot \frac{1}{f(\theta_i) f(\theta_j)} \quad j \geq i \quad \dots \quad (82)$$

Let

$$\left. \begin{aligned} M_1 &= i(\theta_i - \mu_i) - (i-1)(\theta_{i-1} - \mu_{i-1}); \quad i=2, \dots, g-1 \\ M_1 &= (\theta_1 - \mu_1), \quad M_g = -(g-1)(\theta_{g-1} - \mu_{g-1}) \end{aligned} \right\} \dots \quad (83)$$

$$\left. \begin{aligned} M_1^0 &= (g-1)(\theta_i - \mu_i) - (g-i+1)(\theta_{i-1} - \mu_{i-1}); \quad i=2, \dots, g-1 \\ M_1^0 &= (g-1)(\theta_1 - \mu_1), \quad M_g^0 = -(\theta_{g-1} - \mu_{g-1}) \end{aligned} \right\} \dots \quad (84)$$

$$\left. \begin{aligned} N_i &= i(\lambda_i - \gamma_i) - (i-1)(\lambda_{i-1} - \gamma_{i-1}); \quad i=2, \dots, g-1 \\ N_1 &= (\lambda_1 - \gamma_1), \quad N_g = -(\lambda_{g-1} - \gamma_{g-1}) \end{aligned} \right\} \dots \quad (85)$$

$$\left. \begin{aligned} N_1^0 &= (g-1)(\lambda_1 - \nu_1) - (g-i+1)(\lambda_{i-1} - \nu_i); \quad i = 2, \dots, g-1 \\ N_1^0 &= (g-1)(\lambda_1 - \nu_1), \quad N_g^0 = -(\lambda_{g-1} - \nu_g) \end{aligned} \right\} \quad (86)$$

The following can be easily verified.

$$\left. \begin{aligned} (\Sigma + A_1' K A_1)_{ij} &= \frac{1}{g} M_i M_j^0 \quad j > i \\ &= \frac{1}{g} M_j M_i^0 \quad i > j \\ &= \sigma_i^2 + \frac{1}{g} M_i M_i^0 + (\theta_i - \mu_i)(\theta_{i-1} - \mu_i); \quad i=j; i \neq 1, g \\ &= \sigma_i^2 + \frac{1}{g} M_i M_i^0 \quad i = j = 1 \\ &= \sigma_i^2 + \frac{1}{g} M_i M_i^0 \quad i = j = g \end{aligned} \right\} \quad (87)$$

$$\left. \begin{aligned} (T + A_2' K A_2)_{ij} &= -\frac{1}{g} N_i N_j^0 \quad j > i \\ &= -\frac{1}{g} N_j N_i^0 \quad i > j \\ &= \tau_i^2 + \frac{1}{g} N_i N_i^0 + (\lambda_i - \nu_i)(\lambda_{i-1} - \nu_i); \quad i=j; i \neq 1, g \\ &= \tau_i^2 + \frac{1}{g} N_i N_i^0 \quad i = j = 1 \\ &= \tau_i^2 + \frac{1}{g} N_i N_i^0 \quad i = j = g \end{aligned} \right\} \quad (88)$$

$$\left. \begin{aligned} N_1^0 &= (g-1)(\lambda_1 - \nu_1) - (g-i+1)(\lambda_{i-1} - \nu_i); \quad i = 2, \dots, g-1 \\ N_1^0 &= (g-1)(\lambda_1 - \nu_1), \quad N_g^0 = -(\lambda_{g-1} - \nu_g) \end{aligned} \right\} \quad (86)$$

The following can be easily verified.

$$\left. \begin{aligned} (\Sigma + A_1' K A_1)_{ij} &= \frac{1}{g} M_i M_j^0 & j > i \\ &= \frac{1}{g} M_j M_i^0 & i > j \\ &= \sigma_i^2 + \frac{1}{g} M_i M_i^0 + (\theta_i - \mu_i)(\theta_{i-1} - \mu_i); \quad i=j; \quad i \neq 1, g \\ &= \sigma_i^2 + \frac{1}{g} M_i M_i^0 & i = j = 1 \\ &= \sigma_i^2 + \frac{1}{g} M_i M_i^0 & i = j = g \end{aligned} \right\} \quad (87)$$

$$\left. \begin{aligned} (T + A_2' K A_2)_{ij} &= -\frac{1}{g} N_i N_j^0 & j > i \\ &= -\frac{1}{g} N_j N_i^0 & i > j \\ &= \tau_i^2 + \frac{1}{g} N_i N_i^0 + (\lambda_i - \nu_i)(\lambda_{i-1} - \nu_i); \quad i=j; \quad i \neq 1, g \\ &= \tau_i^2 + \frac{1}{g} N_i N_i^0 & i = j = 1 \\ &= \tau_i^2 + \frac{1}{g} N_i N_i^0 & i = j = g \end{aligned} \right\} \quad (88)$$

$$\begin{aligned}
 (E + A_1' K A_2)_{ij} &= \frac{1}{g} M_i N_j^0 & j > i \\
 &= \frac{1}{g} M_i M_j^0 & j < i \\
 &= \vartheta_i \sigma_i \zeta_i + \frac{1}{g} M_i N_i^0 + (\theta_i - \mu_i)(\lambda_{i-1} - \nu_i); & i=j; i \neq 1, g \\
 &= \vartheta_i \sigma_i \zeta_i + \frac{1}{g} M_i N_i^0 & i = j - 1 \\
 &= \vartheta_g \sigma_g \zeta_g + \frac{1}{g} M_g N_g^0 & i = j = g
 \end{aligned} \quad (89)$$

4.4.11. Corollary. The sequence of random variables $\mathcal{Y}(n)$ converges weakly to the random variable \mathcal{Y} which has a multivariate normal distribution with mean vector 0 and variance covariance matrix Λ where

$$\Lambda = T + A_2' K A_2 \quad \dots \quad \dots \quad (90)$$

4.4.12. Corollary. For theorem 4.7 and corollaries 4.8, 4.9, 4.10 and 4.11 to hold good, condition (52) can be replaced by the condition:

For some $\varepsilon > 0$,

$g(y, x)$ is continuous in the regions

$$(\theta_r - \varepsilon < x < \theta_r + \varepsilon, -\infty < y < \infty), \quad r=1, \dots, g-1$$

and $f(\theta_r) \neq 0, \quad r = 1, \dots, g-1.$

(91)

4.4.13. Theorem. Let $S_{(1)} \dots S_{(k)}$ be k independent samples, each of size n , on (Y, X) . Let S be the pooled sample. Let $(\xi_{(1)}(n), \eta_{(1)}(n)), \dots, (\xi_{(k)}(n), \eta_{(k)}(n))$ and $(\xi(n), \eta(n))$ be statistics computed from these samples. Let

$$(\xi^{\circ}(n), \eta^{\circ}(n)) = \frac{1}{\sqrt{k}} \sum_{i=1}^k (\xi_{(i)}(n), \eta_{(i)}(n)) \dots \quad (92)$$

Let condition (91) hold.

Then $(\xi(n), \eta(n)) - (\xi^{\circ}(n), \eta^{\circ}(n)) \rightarrow 0$ in probability.

Proof: The proof of this theorem follows from simple algebra. To illustrate the method of proof we take the simplest case with $k = 2, g = 2$. We should show that

$$\frac{1}{\sqrt{2}} \{(\xi_{(1)}(n), \eta_{(1)}(n)) + (\xi_{(2)}(n), \eta_{(2)}(n))\} - \{\xi(n), \eta(n)\} \\ \text{tends to } 0 \text{ in probability} \dots \dots (93)$$

To show this, it is enough to demonstrate that

$$E \left\{ \frac{1}{\sqrt{2}} \eta_{(1)}(n) + \frac{1}{\sqrt{2}} \eta_{(2)}(n) - \eta(n) \right\}^2 \rightarrow 0 \dots \quad (94)$$

as $n \rightarrow \infty$.

Three other such terms will tend to 0 in a similar way.

Now let us restate our problem.

We have samples S_1 and S_2 , namely

$$(y_{1(1)}, x_{1(1)}), \dots (y_{n(1)}, x_{n(1)})$$

and $(y_{1(2)}, x_{1(2)}), \dots (y_{n(2)}, x_{n(2)})$

on the same random variable (Y, X) . The pooled sample

$$(y_1, x_1) \dots (y_{2n}, x_{2n}) \text{ is denoted by } S.$$

$$\begin{aligned} \eta_{(1)}(n) &= \frac{1}{\sqrt{2}} \eta_{(1)}(n) - \frac{1}{\sqrt{2}} \eta_{(2)}(n) \\ &= \sqrt{n} \left(\sqrt{2} v_1 - \frac{v_{1(1)} + v_{1(2)}}{\sqrt{2}} \right) \\ &= \sqrt{2n} \left(v_1 - \frac{v_{1(1)} + v_{1(2)}}{2} \right) \\ &= \alpha(n) \quad \text{say.} \quad \dots \quad \dots \end{aligned} \quad (95)$$

Now, if $x_{(m+1)(1)} \leq x_{(m+1)(2)}$, then

$$\alpha(n) = \frac{\sqrt{2m}}{2m} \left(\sum_1 y_{(r)(1)} - \sum_2 y_{(r)(2)} \right) \dots \quad (96)$$

where \sum_1 is the summation over all indices r such that

$x_{(m+1)(1)} \leq x_{(r)(1)} < x_{(2m+1)}$ and \sum_2 is the summation over all indices r such that $x_{(2m+1)} \leq x_{(r)(2)} < x_{(m+1)(2)}$. We have another

similar expression when $x_{(m+1)(2)} < x_{(m+1)(1)}$. The probability of

either event is $\frac{1}{2}$. Hence it will suffice to show that

$$E \left[\frac{1}{\sqrt{2n}} \left(\sum_1 y_{(r)}(1) - \sum_2 y_{(r)}(2) \right)^2 \mid \zeta_{(1)} < \zeta_{(2)} \right]$$

tends to 0 as $n \rightarrow \infty \quad \dots \quad \dots \quad (97)$

Let s be the number of observations included in the summation \sum_1 . s is a random variable and the number of observations included in summation \sum_2 is also s . Let us denote $\sqrt{n} (x_{(2s+1)} - \theta_1)$ by ζ_1° .

We can easily verify that

$$E \left[\frac{1}{\sqrt{n}} \sum_1 y_{(r)}(1) \mid \zeta_{(1)}, \zeta_{(2)}, \zeta_1^\circ, s, \zeta_{(1)} < \zeta_{(2)} \right]$$

$$= \frac{s}{\sqrt{n}} \lambda(\theta_1) [1 + o(1)] \quad \dots \quad \dots \quad (98)$$

$$E \left[\left(\frac{1}{\sqrt{n}} \sum_1 y_{(r)}(1) \right)^2 \mid \zeta_{(1)}, \zeta_{(2)}, \zeta_1^\circ, s, \zeta_{(1)} < \zeta_{(2)} \right]$$

$$= \left[\frac{s}{n} \times \text{const.} + \frac{s(s-1)}{n} \lambda^2(\theta_1) \right] [1 + o(1)] \quad \dots \quad (99)$$

Thus

$$E \left[\frac{1}{n} \left(\sum_1 y_{(r)}(1) - \sum_2 y_{(r)}(2) \right)^2 \mid \zeta_{(1)}, \zeta_{(2)}, \zeta_1^\circ, s, \zeta_{(1)} < \zeta_{(2)} \right]$$

$$= \left[\frac{2s}{n} \times \text{const.} - \frac{2s}{n} \lambda^2(\theta_1) \right] [1 + o(1)] \quad \dots \quad (100)$$

$$E \left[\frac{s}{n} \mid \zeta_{(1)}, \zeta_{(2)}, \zeta_1^\circ, \zeta_{(1)} < \zeta_{(2)} \right]$$

$$= \frac{\zeta_1^\circ - \zeta_{(1)}}{\sqrt{n}} f(\theta_1) [1 + o(1)] \quad \dots \quad (101)$$

$$E [| \zeta_1^0 - \zeta_{(1)} | \mid \zeta_{(1)} < \zeta_{(2)}] < \infty \dots \dots \quad (102)$$

(100), (101) and (102) prove the truth of (97). Hence the theorem.

We remark that we can modify the theorem suitably when the samples S_1, \dots, S_k are of varying sizes.

4.5. Theoretical applications.

In this section we show that the limiting distributions of section 4 in effect reduce a large sample S from a population P to just one observation (\bar{y} or \underline{y}) from a multivariate normal distribution. We then ^{derive} the limiting distribution of the measures of divergence $\bar{D}, \bar{\Delta}, \bar{\Gamma}, A, D, \Delta$ and Γ (see section 2) used in fixed interval analysis and in fractile analysis. We indicate the tests that can be used when several (interpenetrating) samples are available from each population. The asymptotic distributions of the concentration ratios (39) and (41) are shown to be normal.

Let samples S and S' be drawn from the populations P and P' respectively, and the method of fixed interval analysis be employed. Corollary 4.5 states that \bar{y} is asymptotically normal with mean \bar{y} and variance covariance matrix $\bar{\Lambda} / n$. We write this in symbols as

$$\bar{y} \rightsquigarrow MN (\bar{y} , \bar{\Lambda} / n) \dots \dots \quad (103)$$

Similarly

$$\bar{v}' \sim MN(\bar{z}', \bar{\Lambda}'/n') \dots \dots \quad (104)$$

(103) and (104) show that the samples S and S' are now reduced to the vectors \bar{v} and \bar{v}' respectively, with asymptotic multivariate normal distributions.

Let $n, n' \rightarrow \infty$ in such a way that $n/n' \rightarrow c, 0 < c < \infty$.

Then

$$\bar{v} - \bar{v}' \sim MN(\bar{z} - \bar{z}', (\bar{\Lambda} + c \bar{\Lambda}')/n) \dots \quad (105)$$

From (105) we can easily obtain the asymptotic distributions of \bar{D} , $\bar{\Delta}$ and $\bar{\Gamma}$ under the null hypothesis (9). $\bar{\Delta}$ has a limiting distribution which is the distribution of $\sum \beta_r \chi_r^2$ where $\chi_1^2, \dots, \chi_g^2$ are independent χ^2 with 1 d.f. and

$$\beta_1 = \left(\frac{\bar{\tau}_1^2}{\kappa_1} + c \frac{\bar{\tau}_1'^2}{\kappa_1} \right), \dots, \beta_g = \left(\frac{\bar{\tau}_g^2}{\kappa_g} + c \frac{\bar{\tau}_g'^2}{\kappa_g} \right).$$

For a particular choice of B , $n \bar{\Gamma}$ is equal to

$n \sum_{i=1}^g [(\bar{v}_i - \bar{v}_i')^2 / (\frac{\bar{\tau}_i^2}{\kappa_i} + c \frac{\bar{\tau}_i'^2}{\kappa_i})]$ and has a limiting distribution that is a χ^2 with g degrees of freedom. The limiting distribution of $\sqrt{n} \bar{D}$ exists but does not have a simple algebraic form.

Let $S_{(1)}, \dots, S_{(k)}$ be k independent and equally valid (interpenetrating) sub-samples of the same size n from the population P . Let $S'_{(1)}, \dots, S'_{(k')}$ be k' interpenetrating subsamples of size n' each from P . If n and n' are large these two sets of samples can be reduced to two samples $\bar{v}_{(1)}, \dots, \bar{v}_{(k)}$ and $\bar{v}'_{(1)}, \dots, \bar{v}'_{(k')}$ from multivariate normal distributions with parameters* $(\bar{v}, \bar{\Lambda}/n)$ and $(\bar{v}', \bar{\Lambda}'/n')$, respectively.

The problem of fixed interval analysis is to test the hypothesis $\bar{v} = \bar{v}'$. Since $\bar{\Lambda}$ and $\bar{\Lambda}'$ are diagonal the problem can be viewed as the problem of simultaneous independent Fisher-Behren tests.

To pose the problem as one in classical multivariate analysis we should strengthen hypothesis (9) to the hypothesis

$$\bar{v} = \bar{v}', \quad \bar{\Lambda} = \bar{\Lambda}' \quad \dots \quad (106)$$

Multivariate analysis now yields us two solutions to this problem.

We can use the likelihood ratio criterion

$$\prod_{r=1}^k \left\{ [k s_r^2 + k' s_r'^2 + k k' (\bar{v}_r - \bar{v}'_r)^2] / [k s_r^2 + k' s_r'^2] \right\} \dots (107)$$

where $\bar{v}_r = \sum_{j=1}^k \bar{v}_r(j) / k$

$$s_r^2 = \sum_{j=1}^k (\bar{v}_r(j) - \bar{v}_r)^2 / k \quad \dots \quad (108)$$

* the parameters are the mean vector and the variance covariance matrix, respectively.

Large values of this criterion will form the region of rejection.

The distribution of this criterion has been evaluated by Box (1949).

Another method is to use the criterion

$$\sup_{1 \leq r \leq g} \frac{|\bar{v}_r - \bar{v}'_r| [ckk' (k + k' - 2)]^{\frac{1}{2}}}{[(cks_r^2 + k's_r'^2)(k' + kc)]^{\frac{1}{2}}} \dots \quad (109)$$

The distribution of this criterion is the distribution of the absolute maximum of g independent t -distributions each with $k + k' - 2$ degrees of freedom. Significance points of this distribution can easily be obtained from the tables of the t -distribution.

Let samples S and S' be drawn from the populations P and P' respectively, and the method of fractile analysis be used.

Corollary 4.11 states that

$$\underline{v} \sim MN(\underline{\gamma}, \Lambda / m) \dots \quad (110)$$

$$\underline{v}' \sim MN(\underline{\gamma}', \Lambda' / m') \dots \quad (111)$$

(110) and (111) show that the samples S and S' are now reduced to the vectors \underline{v} and \underline{v}' respectively, with asymptotic multivariate normal distributions.

Let $n, n' \rightarrow \infty$ in such a way that $n/n' \rightarrow c, 0 < c < \infty$.

Then

$$\underline{v} - \underline{v}' \sim MN(\underline{\gamma} - \underline{\gamma}', (\Lambda + c\Lambda')/m) \dots \quad (112)$$

From (112) we can obtain the asymptotic distribution of A , D , Δ and Γ under the null hypothesis (19). $m \Delta$ has a limiting distribution which is the distribution of $\sum \beta_r \chi_r^2$ where $\chi_1^2, \dots, \chi_g^2$ are independent χ^2 with 1 d.f. and β_1, \dots, β_g are the latent roots of $(\Lambda + c \Lambda')$. Γ has a similar limiting distribution. In particular

$$\Gamma \rightarrow m(\underline{y} - \underline{y}')(\Lambda + c \Lambda')^{-1}(\underline{y} - \underline{y}')' \dots \quad (113)$$

has a limiting distribution that is a χ^2 with g d.f. The limiting distribution of $\sqrt{m} A$ and $\sqrt{m} D$ exist but do not have simple algebraic expressions. Crude approximations to the limiting distribution of $\sqrt{m} A$ can be made by the use of the following easily proved inequality

$$\frac{\sqrt{\Delta}}{6} \leq A \leq \sqrt{g} \cdot \sqrt{\Delta} \dots \quad (114)$$

Let $S_{(1)}, \dots, S_{(k)}$ ($S'_{(1)}, \dots, S'_{(k')}$) be $k(k')$ interpenetrating subsamples from $P(P')$. If $n(n')$ is large then

$\underline{y}_{(1)}, \dots, \underline{y}_{(k)}$ ($\underline{y}'_{(1)}, \dots, \underline{y}'_{(k')}$) is a sample from a multivariate normal distribution with parameters $(\underline{y}, \Lambda/m)$ ($(\underline{y}', \Lambda'/m')$).

The problem of fractile analysis is to test the hypothesis $\underline{y} = \underline{y}'$. To tackle this problem as one in multivariate analysis we use a restricted hypothesis

$$\underline{y} = \underline{y}', \quad \Lambda = \Lambda' \quad \dots \quad (115)$$

We now have the familiar problem of testing the equality of the mean when the variance covariance matrix of one multivariate normal population is a constant multiple of ^{that of} the other. The Mahalanobis D^2 -statistic will be used. Let

$$\begin{aligned} \underline{v}_i^0 &= \sum_{\alpha=1}^k \underline{v}_i(\alpha) / k \\ s_{ij} &= \sum_{\alpha=1}^k (\underline{v}_i(\alpha) - \underline{v}_i^0)(\underline{v}_j(\alpha) - \underline{v}_j^0) / k \quad \dots \end{aligned} \quad (116)$$

$$\begin{aligned} s &= (s_{ij}) \\ \bar{s} &= [kcs + k's'] / c \quad \dots \quad \dots \end{aligned} \quad (117)$$

Then

$$\frac{k + k' - g - 1}{g} \cdot \frac{kk'}{ck + k'} (\underline{v}^0 - \underline{v}^0') (\bar{s})^{-1} (\underline{v}^0 - \underline{v}^0') \quad (118)$$

is our test criterion. Its distribution is an F distribution with g and $k + k' - g - 1$ d.f.

Let \underline{y} (\underline{y}') be derived from $S(S')$, the sample obtained by pooling $S_{(1)}, \dots, S_{(k)}$ ($S'_{(1)}, \dots, S'_{(k)}$). As an application of Theorem 4.13 we can substitute \underline{y} and \underline{y}' for \underline{v}^0 and \underline{v}^0' in (118) without changing the limiting distribution.

In the preceding discussion we assumed that the interpenetrating samples from each population are of the same size. We can remove this restriction by simple modifications. The tests mentioned above require several samples from each populations, and make use of them only through the \bar{y}'_s or y'_s . This involves considerable labour and waste of information. When only one sample is available from each population we cannot make use of the measures of divergence Δ , D , etc., for testing since their limiting distributions involve unknown constants. In the next chapter we suggest methods of overcoming this difficulty.

In section 2 we described how concentration curves can be drawn with the help of the data of fixed interval analysis and fractile analysis. We now give the limiting distributions of the concentration ratio, as a direct application of the theorems of section 4 and a theorem of Greiner (1946) pp. 366.

$$\text{Let } \bar{\gamma} = \frac{\sum \pi_i \bar{y}_i (\bar{\pi}_{i-1} + \bar{\pi}_i)}{(\sum \pi_i \bar{y}_i)} - 1 \quad \dots \quad (119)$$

$$\bar{\pi}_i = \pi_1 + \dots + \pi_i \quad \dots \quad (120)$$

$$\frac{(\bar{\pi}_{i-1} + \bar{\pi}_i)}{(\sum \pi_r \bar{y}_r)} - \frac{\bar{\gamma}}{(\sum \pi_r \bar{y}_r)} = \bar{d}_i \quad i = 1, \dots, g \quad (121)$$

$$\frac{2 \bar{y}_i \bar{\pi}_{i-1} + \sum_1^g \pi_r \bar{y}_r}{(\sum \pi_r \bar{y}_r)} = \frac{\bar{\gamma}}{(\sum \pi_r \bar{y}_r)} \cdot \bar{y}_i = e_i \quad i = 1, \dots, g \quad (122)$$

Then $\sqrt{n} (\bar{C} - \bar{\gamma})$ has a limiting distribution that is normal with mean 0 and variance

$$\sum_{i=1}^g d_i^2 \bar{v}_i^2 + e \bar{K} e' \quad \dots \quad \dots \quad (123)$$

Let $\gamma = \frac{2 \sum_{i=1}^g v_i}{g(v_1 + \dots + v_g)} = \frac{k+1}{g} \quad \dots \quad \dots$ (124)

$$\frac{2i}{g(v_1 + \dots + v_g)} = \frac{\bar{\gamma}}{(v_1 + \dots + v_g)} = d_i, \quad i = 1, \dots, g \quad (125)$$

Then $\sqrt{n} (C - \gamma)$ has a limiting distribution that is normal with mean 0 and variance

$$\underline{d} \wedge \underline{d}' \quad \dots \quad \dots \quad (126)$$

Though we have not explicitly mentioned, it should be noted that (46) is assumed when fixed interval analysis is employed and (91) is assumed when fractile analysis is employed, in this section.

4.6. A general formulation for the regression problem.

In this section we present a general formulation of the regression problem. It turns out that the methods of fixed interval analysis and fractile analysis are convenient approximations to this general method.

We propose several measures of divergence between two regressions as in fixed interval analysis and fractile analysis. We do not give the distributions of these measures, even for large samples, since we were not able to obtain them. In the present case we would have to deal with the convergence of not only multivariate normal distributions but of random functions and this proves to be a very difficult task. The methods available at present are purely descriptive.

Let (Y, X) be a random variable.

$$E(Y | X = x) = \lambda(x), \quad -\infty < x < \infty \quad \dots \quad (127)$$

$$E(Y | X \leq x) = \bar{\nu}(x), \quad -\infty < x < \infty \quad \dots \quad (128)$$

$$E(Y | F(X) \leq t) = \nu(t), \quad 0 < t < 1, \quad \dots \quad (129)$$

The problem of testing the equality of regression in two populations (Y, X) and (Y', X') can be formulated as testing some one of the following hypothesis

$$H_1 : \lambda(x) = \lambda'(x) \quad \text{for all } x, \quad -\infty < x < \infty \quad (130)$$

$$H_2 : \bar{\nu}(x) = \bar{\nu}'(x) \quad \text{for all } x, \quad -\infty < x < \infty \quad (131)$$

$$H_3 : \nu(t) = \nu'(t) \quad \text{for all } t, \quad 0 < t < 1 \quad (132)$$

H_1 , H_2 and H_3 correspond, respectively, to the problems considered in classical regression theory, fixed interval analysis and fractile analysis.

We have expressed "the equality of two regression" in three different ways. It would be natural to estimate the quantities involved in a hypothesis and test by using some criterion of divergence between the estimates. For example, the method of testing H_2 ^{maybe} as follows: Let $(y_1, x_1), \dots, (y_n, x_n)$ be a sample S of independent observations on (Y, X) . Let

$$\bar{v}(x) = \sum_{x_1 \leq x} y_1 / n_x, \quad -\infty < x < \infty \quad \dots \quad (133)$$

where n_x is the number of observations in S that are less than or equal to x . $\bar{v}(x)$ is an estimate of $\bar{V}(x)$. Similarly by taking a sample S' on (Y', X') we construct an estimate $\bar{v}'(x)$ of $\bar{V}'(x)$. Consider

$$\bar{R} = \sup_{-\infty < x < \infty} |\bar{v}(x) - \bar{v}'(x)| \quad \dots \quad (134)$$

as the measure of divergence. In order to study the asymptotic distribution of \bar{R} , consider the stochastic process

$$\bar{Z}_n(x) = \sqrt{n} (\bar{v}(x) - \bar{V}(x)) \quad \dots \quad (135).$$

We can easily show, by means of corollary 4.5, that the finite dimensional distributions of $\bar{Z}_n(x)$ converge to multivariate normal distributions. It not known whether the limiting process, which can be constructed from these finite dimensional limiting

distributions, is realized in $C(X)$, the space of all continuous functions on X with the topology of uniform convergence on compacta, not to speak of the convergence of the processes. Even if this were so, it is unlikely that the limiting distribution of $\sqrt{n} \bar{R}$ exists. The seeming lack of success of this method may be due to our choice of the measure of divergence. It would be interesting to develop this method further to see whether it can provide convenient tests.

In trying to estimate the parameters involved in the specification of H_3 we may proceed as follows. The sample S is rearranged with the x 's in increasing order of magnitude as follows:

$$(y_{(1)}, x_{(1)}), \quad \dots \quad (y_{(n)}, x_{(n)}) \quad \dots \quad (136)$$

Let

$$v(t) = \sum_{r \leq [nt]} y_r / [nt] \quad 0 < t \leq 1 \quad \dots \quad (137)$$

$v'(t)$ is defined in the same way. We may consider several measures of divergence.

$$R = \sup_{0 < t < 1} | v(t) - v'(t) | \quad \dots \quad (138)$$

$$D = \int_0^1 | v(t) - v'(t) | dt \quad \dots \quad (139)$$

$$\Delta = \int_0^1 (v(t) - v'(t))^2 dt \quad \dots \quad (140)$$

Using corollary 4.11 we can show that the finite dimensional distributions of the process

$$Z_n(t) = \sqrt{n} (v(t) - \nu(t)) \dots \quad (141)$$

converge. Here also we are able to ascertain neither the realization of the limiting process in $C(0, 1)$, $L_1(0, 1)$ or $L_2(0, 1)$, nor the convergence of the processes.

The above corresponds to fractile analysis with $g = n$ and $m = 1$. In the theorems of section 4 we assumed that $m \rightarrow \infty$.

It is therefore plausible that if $n = g_n \cdot m_n$ and both g_n and $m_n \rightarrow \infty$ suitably, R , D and Δ would have asymptotic distributions.

It is also plausible that an estimate $\hat{\nu}(t)$ of $\nu(t)$ of the following sort

$$\begin{aligned} \hat{\nu}(t) &= 0 && \text{if } t < 1/g_n \\ &= v(1/g_n) && \text{if } 1/g_n \leq t < 2/g_n \quad \dots \quad (142) \\ &\dots && \dots \\ &= v\left(\frac{g_n - 1}{g_n}\right) && \text{if } \frac{g_n - 1}{g_n} \leq t < 1 \end{aligned}$$

would be useful. Such a technique has been used successfully by Parthasarathy and Bhattacharya (1961) in connection with testing for a specified regression function $\mu(x)$.

Let

$$\begin{aligned} \tilde{y}(t) &= \frac{y(1) + \dots + y(m_n)}{m_n} & 0 \leq t \leq \frac{1}{\varepsilon_n} \\ &= \frac{y(m_n+1) + \dots + y(2m_n)}{m_n} & \frac{1}{\varepsilon_n} < t \leq \frac{2}{\varepsilon_n} \quad \dots \quad (143) \\ &\dots \dots \dots \dots \dots \dots \dots \\ &= \frac{y(\varepsilon_n^{-1} - m_n + 1) + \dots + y(\varepsilon_n m_n)}{m_n} & \frac{\varepsilon_n - 1}{\varepsilon_n} < t \leq 1 \end{aligned}$$

$$\begin{aligned} \tilde{\mu}(t) &= \frac{\mu(1) + \dots + \mu(m_n)}{m_n} & 0 \leq t \leq 1/\varepsilon_n \\ &= \frac{\mu(m_n+1) + \dots + \mu(2m_n)}{m_n} & \frac{1}{\varepsilon_n} < t \leq \frac{2}{\varepsilon_n} \quad (144) \\ &\dots \dots \dots \dots \dots \dots \dots \\ &= \frac{\mu(\varepsilon_n^{-1} - m_n + 1) + \dots + \mu(\varepsilon_n m_n)}{m_n} & \frac{\varepsilon_n - 1}{\varepsilon_n} < t \leq 1 \end{aligned}$$

where

$$\mu(r) = \mu(x(r)) \dots \dots \quad (145)$$

Parthasarathy and Bhattacharya (1961) evaluate the asymptotic distribution of

$$\tilde{R}_n = \sup_{1 \leq r \leq \varepsilon_n} \frac{\left| \tilde{y}\left(\frac{r}{\varepsilon_n}\right) - \tilde{\mu}\left(\frac{r}{\varepsilon_n}\right) \right|}{\sqrt{\tilde{\beta}_r / m_n}} \dots \quad (146)$$

where the $\tilde{\beta}_r$'s depend on the conditional variance of Y given X.

$$\text{Let } \beta_m(x) = E(|Y - \mu(x)|^m | X = x) \quad (147)$$

$$\text{Let } \beta_p(x) \text{ exist for all } x \text{ for some } p > 2. \quad \dots \quad (148)$$

$$\beta_2(x) > c_1 > 0 \text{ for all } x \quad \dots \quad (149)$$

$$\beta_m(x) < c_2 < \infty \text{ for all } x \text{ and } m = 1, \dots, p. \quad \dots \quad (150)$$

$$\mu(x) \text{ be continuous in } x \quad \dots \quad (151)$$

$$X \text{ have a strictly increasing continuous distribution} \\ \text{function} \quad \dots \quad (152)$$

$$m_n \geq (\varepsilon_n \log \varepsilon_n)^{2/p-2} \quad \dots \quad (153)$$

4.6.1. Theorem (Parthasarathy and Bhattacharya).

Under the conditions (148), (149), (150), (151), (152) and (153)

$$\text{Prob} \left\{ \frac{\tilde{R}_n^2 - 2 \log \varepsilon_n + \log \log \varepsilon_n}{2} < x \right\} \\ \rightarrow \exp\left(-\frac{1}{\sqrt{\pi}} e^{-x}\right) \quad \dots \quad (154)$$

as $\varepsilon_n \rightarrow \infty$.

Parthasarathy and Bhattacharya have also suggested methods of estimating $\tilde{\beta}_r$'s but they are too cumbersome to be stated here.

Their method cannot be modified to test the equality of two regressions.

It may be hoped, however, that by letting both ε_n and m_n tend to infinity suitably we can evaluate the limiting distributions of the

statistics \bar{R} , R , Δ and D .

In conclusion, it may ^{added that} be Δ the results of sections 4,5 remain useful in practice. In many instances a fixed g is more meaningful than a g tending to infinity. There is also the practical difficulty of evaluating the measures of divergence ^{of (134), (138), (139) and (140), in} \bar{R} , R , Δ and D Δ large samples.

Chapter V.

FIXED INTERVAL ANALYSIS AND FRACTILE ANALYSIS - PART - II

5.1. Introduction and summary.

In the last chapter we used the limit theorems of section 4.4 to reduce a sample S from a population P to a single observation on a multivariate normal distribution (See section 4.5). This fact enables us to use classical multivariate analysis in fixed interval analysis and fractile analysis, and forces us to collect several samples from each of the two populations. We remarked that such tests therefore necessitate the considerable labour of collecting several interpenetrating subsamples without, at the same time, making full use of the samples or the special nature, if known, of the distribution of (Y, X) . There is thus room for development for more efficient tests. We proceed to construct such tests in this chapter.

Measures of divergence, that could be used for testing purposes were defined in section 4.2. These measures $\bar{D}, \bar{\Delta}, \bar{\Gamma}, A, D, \Delta, \Gamma$ etc., can be calculated even if just one sample is available from each of the two populations P and P' . The limiting distributions of these criteria of divergence have been derived in section 4.5. These limiting distributions involve unknown population constants. In this chapter we construct consistent estimates for these

unknown constants when just one sample is available from each population, and show that the limiting distributions are unaltered by the substitution of these estimates.

For the case of fixed interval analysis we exhibit in section 2 a method of constructing such consistent estimates without any assumptions on the distribution of (Y, X) . This allows us to construct nonparametric tests for the problem of fixed interval analysis.

For the problem of fractile analysis we are not able to give methods of obtaining consistent estimates in the general case. We find (in section 3), however, that the asymptotic variance covariance matrix of \hat{y} assumes a simple form when the vector (Y, X) is normally distributed. This fact is used to estimate this matrix consistently. Tables have been provided for this purpose. The other important case, wherein (Y, X) has a bivariate lognormal distribution, is also dealt with.

Some model sampling experiments were conducted to investigate the rate at which the various limiting distributions (of section 4.4) are approached. The results are given in sections 4 and 5. These results are purely empirical, and theoretical investigations of the problem remain outstanding.

The problem of formulating alternatives to the null hypotheses of fixed interval analysis and fractile analysis remains to be inves-

investigated. Questions of power and efficiency of our tests should be considered. Satisfactory test procedures are still to be developed in small samples and in designs with multistage sampling.

The notations and definitions to be used in this chapter are the same as in Chapter IV. It will be assumed that in any discussion on fixed interval analysis (46) of Chapter IV holds. In the same way, (91) of Chapter IV will be assumed to hold in discussion on fractile analysis.

5.2. Testing in fixed interval analysis.

Among the statistics \bar{D} , $\bar{\Delta}$ and \bar{F} (of (14), (15) and (16) respectively, of Chapter IV) that can be used to test the null hypothesis ((9) of Chapter IV) of fixed interval analysis, \bar{F} is most suited for practical use since its limiting distribution (see section 4.5) is the most simple. Further \bar{F} corresponds to the statistic that is used for testing simultaneous that several means are zero when all the variances are known. When the variables concerned are independently normally distributed this test corresponds to the Hotelling's T. Thus \bar{F} seems to be the best criterion we can use in this situation.

$$\bar{F} = \sum [(\bar{v}_1 - \bar{v}'_1)^2 / (\frac{\bar{\tau}_i^2}{\kappa_1 n} + \frac{\bar{\tau}'_i{}^2}{\kappa_1 n'})] \dots \quad (1)$$

We note that $\frac{\bar{\tau}_i^2}{n_1}$ and $\frac{\bar{\tau}'_i{}^2}{n'_1}$ are consistent for $\frac{\tau_i^2}{\kappa_1 n}$ and

and $\frac{\bar{c}_i'^2}{n_1 n_1'}$, $i = 1, \dots, g$, the unknown constants that enter in \bar{F} .

Let

$$\bar{F}^* = \sum [(\bar{v}_1 - \bar{v}_1')^2 / (\frac{\bar{c}_1^2}{n_1} + \frac{\bar{c}_1'^2}{n_1'})] \dots \quad (2)$$

We note that \bar{F}^* can be calculated from S and S' and that its limiting distribution is a χ^2 with g degrees of freedom. The critical region for testing the null hypothesis of fixed interval analysis will be the region of large values of \bar{F}^* .

Our applications so far (of the Theorems 4.4.4, 4.4.5 etc.) depend in essence on the fact that in large samples we can treat $\bar{v}_1, \dots, \bar{v}_g$ as independently normally distributed variables. This can be used in many ways. For instance, if we have two samples $S_{(1)}$ and $S_{(2)}$ (pooled, they form S) from one population P and only one, S' , from the second population P' , our test criterion would be

$$\bar{F}^* = \sum [(\bar{v}_1 - \bar{v}_1')^2 / (\frac{n_{1(1)} \bar{c}_1^2(1) + n_{1(2)} \bar{c}_1^2(2)}{n_{1(1)} \cdot n_{1(2)}} + \frac{\bar{c}_1'^2}{n_1'})] \dots (2a)$$

The limiting distribution of this statistic is again χ^2 with g d.f.

5.3. Testing in fractile analysis.

We have seen in section 4.5. (see (110) of Chapter IV) that

$$\underline{v} \sim MN(\underline{\gamma}, \Lambda/m) \dots \quad (3)$$

Let $\lambda_{ij} = (\wedge)_{ij}$. Then from (88) of Chapter IV we have

$$\left. \begin{aligned} \lambda_{ij} = \lambda_{ji} &= \frac{1}{g} N_i N_i^0 & j > i \\ &= \gamma_i^2 + \frac{1}{g} N_i N_i^0 + (\lambda_{i-1} - \gamma_i)(\lambda_i - \gamma_i); \quad i=j; \quad i \neq 1, g \\ &= \gamma_i^2 + \frac{1}{g} N_i N_i^0 & i = j = 1 \\ &= \gamma_g^2 + \frac{1}{g} N_g N_g^0 & i = j = g \end{aligned} \right\} \quad (4)$$

where

$$\left. \begin{aligned} N_i &= i(\lambda_i - \gamma_i) - (i-1)(\lambda_{i-1} - \gamma_i); \quad i=2, \dots, g-1 \\ N_1 &= (\lambda_1 - \gamma_1), \quad N_g = - (g-1)(\lambda_{g-1} - \gamma_g) \end{aligned} \right\} \dots \quad (5)$$

$$\left. \begin{aligned} N_i^0 &= (g-i)(\lambda_i - \gamma_i) - (g-i+1)(\lambda_{i-1} - \gamma_i) \\ N_1^0 &= (g-1)(\lambda_1 - \gamma_1), \quad N_g^0 = - (\lambda_{g-1} - \gamma_g) \end{aligned} \right\} \dots \quad (6)$$

There is an important difference between fixed interval analysis and fractile analysis. Relations (103) and (110) of Chapter IV show that $\underline{\bar{v}}$ and \underline{v} are asymptotically normal with variance

covariance matrix $\bar{\Lambda}/n$ and Λ/m , respectively. $\bar{\Lambda}$ is always diagonal, but Λ is not diagonal in general. Λ is diagonal only when $\lambda_i = \nu_i = \lambda_{i-1}$ for all i . This does not hold in general, for instance when $\lambda(x)$ is strictly monotone. Thus whereas $\bar{v}_1, \dots, \bar{v}_g$ can be considered to be independent in large samples, v_1, \dots, v_g are dependent for all n .

Among the several measures of divergence introduced in section 4.2, Δ and Γ can be used since their limiting distributions under the null hypothesis is of a simple form.

$$n \Delta = \sum m (v_i - v'_i)^2 \quad \dots \quad \dots \quad (7)$$

has a limiting distribution which is that of $\sum_1^g \beta_i \chi_i^2$ where $\chi_1^2, \dots, \chi_g^2$ are independent chisquares with 1 d.f. and β_1, \dots, β_g are the latent roots of $\Lambda + c \Lambda'$. This distribution can be approximated by [See Satterthwaite (1946). For other approximations see Robbins and Pitman (1949)]. dZ where Z has a χ^2 -distribution with α d.f.; d and α are given by the relations

$$d = \sum \beta_i^2 / \sum \beta_i \quad \dots \quad \dots \quad (8)$$

$$\alpha = (\sum \beta_i)^2 / \sum \beta_i^2 \quad \dots \quad \dots \quad (9)$$

$$\Gamma = (\underline{v} - \underline{v}') \left(\frac{\Lambda}{n} + \frac{\Lambda'}{m'} \right)^{-1} (\underline{v} - \underline{v}')' \quad \dots \quad \dots \quad (10)$$

has a limiting distribution that is a χ^2 with g d.f.

These two statistics cannot be used in practice unless Λ and Λ' are known. Let us try to estimate Λ from the sample S . It seems natural to use t_1^2 as an estimate of λ_{11} . But our extensions of Hoeffding's Theorems on order statistics in section 3.7 (relation (113) of Chapter III) show that

$$E(t_1^2) \rightarrow \tau_i^2 \text{ as } m \rightarrow \infty. \quad \dots \quad (11)$$

We can show by assuming some extra moments for Y , that the variance of t_1^2 tends to zero. Thus t_1^2 would be consistent for τ_i^2 . From the expression for λ_{11} in (4) we see that τ_i^2 is equal to λ_{11} if and only if $\lambda_i = \nu_i = \lambda_{i-1}$, and this imposes a condition on the distribution of (Y, X) . Thus t_1^2 is not consistent for λ_{11} , in general. This fact seems to be a great set back in the construction of nonparametric tests in the method of fractile analysis.

We therefore proceed to construct convenient test procedures when (Y, X) is known to follow some special distributions.

Let us assume that the distribution of (Y, X) is bivariate normal with parameters $\nu, \mu; \tau^2, \sigma^2; \rho$. We now evaluate the matrix for this distribution. We note that Λ does not depend on ν, μ and σ^2 . Using (4) we then find that Λ can be written down in the form

$$\Lambda = \tau^2 I + \rho^2 \tau^2 Q \quad \dots \quad (12)$$

where I is the identity matrix and Q is a $g \times g$ square matrix depending only on g . Let

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2) \quad \dots \quad (13)$$

$$\theta_i \text{ is the solution of } \int_{-\infty}^{\theta} \varphi(x) dx = i/g; \quad i=0, 1, \dots, g-1 \quad \dots (14)$$

$$\mu_i = g(\varphi(\theta_{i-1}) - \varphi(\theta_i)) \quad i = 1, \dots, g \quad (15)$$

$$\left. \begin{aligned} d_i &= g(\theta_{i-1} \varphi(\theta_{i-1}) - \theta_i \varphi(\theta_i)) \quad i=2, \dots, g-1 \\ d_1 &= -g \theta_1 \varphi(\theta_1), \quad d_g = g \theta_{g-1} \varphi(\theta_{g-1}) \end{aligned} \right\} (16)$$

$$\left. \begin{aligned} M_i &= i(\theta_i - \mu_i) - (i-1)(\theta_{i-1} - \mu_i); \quad i=2, \dots, g-1 \\ M_1 &= (\theta_1 - \mu_1), \quad M_g = -(g-1)(\theta_{g-1} - \mu_g) \end{aligned} \right\} (17)$$

$$\left. \begin{aligned} M_i^0 &= (g-1)(\theta_i - \mu_i) - (g-i+1)(\theta_{i-1} - \mu_i); \quad i=2, \dots, g-1 \\ M_1^0 &= (g-1)(\theta_1 - \mu_1), \quad M_g^0 = -(\theta_{g-1} - \mu_g) \end{aligned} \right\} (18)$$

The elements of the matrix Q are given by

$$\left. \begin{aligned} (Q)_{ij} &= (Q)_{ji} = \frac{1}{g} M_i M_j^0 \quad \text{if } j > i \\ &= (d_i - \mu_i^2) + \frac{1}{g} M_i M_i^0 + (\theta_{i-1} - \mu_i)(\theta_i - \mu_i); \quad i=j; i \neq 1, g \\ &= d_i - \mu_i^2 + \frac{1}{g} M_i M_i^0 \quad \text{if } i = j = 1 \\ &= d_g - \mu_g^2 + \frac{1}{g} M_g M_g^0 \quad \text{if } i = j = g \end{aligned} \right\} (19)$$

It can be easily demonstrated that the matrix Q is doubly symmetric, i.e.

$$(Q)_{ij} = (Q)_{ji} = (Q)_{g-i+1, g-j+1} = (Q)_{g-j+1, g-i+1} \quad (20)$$

The matrix Q has certain interesting properties (not basic to our main work) which are easily derived from the fact that $\frac{1}{g} \sum v_i = \bar{y}$, the mean of y 's. Thus, Q is negative semi-definite, $\sum_j (Q)_{ij} = 0$ for each i , one of the latent roots of Q is zero and so on.

In table I, at the end of this chapter, we have given these matrices Q when $g = 2(1) 16, 20, 25$. Since the matrix Q is doubly symmetric, we give only $(Q)_{ij}$ for $j \leq i \leq g-j+1$; $1 \leq j \leq [\frac{g+1}{2}]$ for each g .

We have also tabulated the latent roots (q_1, \dots, q_g) of Q in table II for $g = 2(1) 16, 20, 25$. Actually Q, q_1, μ_1 etc. all depend on g , so that $Q_g, q_{g,i}, \mu_{g,i}$ etc., would be the more appropriate notations for them. We however drop the suffix g whenever we feel that it will not cause confusion.

Let $\hat{\tau}^2$ and $\hat{\rho}^2 \tau^2$ be consistent estimates of τ^2 and $\rho^2 \tau^2$ respectively. Λ is now consistently estimated by $\hat{\Lambda} = \hat{\tau}^2 I + \hat{\rho}^2 \tau^2 Q$. (Q can be got from Table I for some values of g .) Then

$$\Gamma^* = (\underline{y} - \underline{y}') \left(\frac{\hat{\Lambda}}{m} + \frac{\hat{\Lambda}'}{m'} \right)^{-1} (\underline{y} - \underline{y}') \dots \quad (21)$$

will be distributed as a χ^2 with g d.f.

The latent roots of $\hat{\Lambda}$ are given by $\hat{\tau}_1^2 + \hat{\rho}^2 \tau_1^2 q_1, \dots, \hat{\tau}_g^2 + \hat{\rho}^2 \tau_g^2 q_g$

and these are consistent estimates of the latent roots of Λ .

Let us denote the consistent ^{estimates of the} latent roots of $\Lambda + c \Lambda'$ (which is got in the same way) by $\hat{\beta}_1, \dots, \hat{\beta}_g$. Then the limiting distribution of $n \Delta$ can be approximated by $\sum \hat{\beta}_i \chi_i^2$ where $\chi_1^2, \dots, \chi_g^2$ are independent chisquares with 1 d.f. We should note that large values of the statistics $n \Delta$ and Γ^* form the critical region for testing the null hypothesis ((19) of Chapter IV) of fractile analysis.

We now suggest two methods of estimating τ^2 and $g^2 \tau^2$ from the sample S . Let t^2 and r be the sample variance of y and the sample correlation coefficient between y and x . Then t^2 and $r^2 t^2$ are consistent estimates of τ^2 and $g^2 \tau^2$ respectively. These can now be used as suggested in the preceding paragraph. Since the sample size of S will be large it will be time consuming to compute t^2 and r . We now suggest another method that will be computationally easier.

This method makes use of the sample S only through u_i 's and v_i 's. Consider

$$\hat{t}^2 = \frac{1}{g} \sum v_i^2 - \left(\frac{1}{g} \sum v_i \right)^2 \dots \quad (22)$$

$$\tilde{s}^2 = \frac{1}{g} \sum u_i^2 - \left(\frac{1}{g} \sum u_i \right)^2 \quad \dots \quad (23)$$

$$\tilde{r} \tilde{s} \tilde{t} = \frac{1}{g} \sum u_i v_i - \left(\frac{1}{g} \sum u_i \right) \left(\frac{1}{g} \sum v_i \right) \quad \dots \quad (24)$$

We can easily demonstrate (by showing that the expectations converge and variances tend to zero) that $m \tilde{t}^2$, $m \tilde{s}^2$ and $m \tilde{r} \tilde{s} \tilde{t}$ are consistent for

$$\frac{g-1}{g} \sigma^2 + \frac{1}{g} \sigma^2 \rho^2, \quad \frac{g-1}{g} \sigma^2 + \frac{1}{g} \sigma^2 L \quad \text{and} \quad \frac{g-1}{g} \rho \sigma \tau + \frac{1}{g} \rho \sigma \tau L$$

respectively, where

$$L = T_Y Q + \mu_{g,1}^2 + \dots + \mu_{g,g}^2 \quad \dots \quad (25)$$

$$T_Y Q = q_{g,1} + \dots + q_{g,g} \quad \dots \quad (26)$$

In tables III and IV we have given $T_Y Q_g$ and $\sum_1^g \mu_{g,i}^2$ for $g = 2(1)16, 20, 25$. Making use of these tables, we can easily obtain consistent estimates for σ^2 and $\rho^2 \sigma^2$ based on \tilde{t} , \tilde{s} and \tilde{r} .

A distribution of (Y, X) which is very important in econometrics is the bivariate lognormal distribution. It is known that the distribution of expenditure on an item, food, say, and the total expenditure is often closely approximated by the bivariate lognormal. See Roy and Dhar (1959), Iyengar (1960).

The case of the bivariate lognormal can be reduced to that of a bivariate normal distribution by taking the logarithms of both the variables. This method is time consuming and not practicable in large samples.

Let us now inspect the form of the matrix Λ when the distribution of $(\log Y, \log X)$ is bivariate normal with parameters $\nu, \mu, \tau^2, \sigma^2, \rho$.

Let $\varphi(x)$ and $\Phi(x)$ be defined by (13) and (14) respectively.

$$\text{Let } \Phi(x) = \int_{-\infty}^x \varphi(t) dt \quad \dots \quad \dots \quad (27)$$

$$\lambda_1 = \left[\exp \left(\nu + \frac{\tau^2}{2} \right) \right] \varphi(\theta_1 - \rho\tau) / \varphi(\theta_1) \quad \dots \quad (28)$$

$$\lambda_2 = \left[\exp \left(\nu + \frac{\tau^2}{2} \right) \right] \cdot \sigma \cdot \left[\Phi(\theta_1 - \rho\tau) - \Phi(\theta_{1-1} - \rho\tau) \right] \quad (29)$$

$$\tau_i^2 = \left[\exp \left(2\nu + \tau^2 \right) \right]$$

$$\times \left[\sigma^2 \left(\Phi(\theta_1 - \rho\tau) - \Phi(\theta_{1-1} - \rho\tau) \right) - \sigma^2 \left(\Phi(\theta_1 - \rho\tau) - \Phi(\theta_{1-1} - \rho\tau) \right)^2 \right]$$

$$\dots \quad (30)$$

We can now find λ_{1j} from (4) after calculating N_1, N_1^0 from (5).

We note that $\exp \left(\nu + \frac{\tau^2}{2} \right)$ is the mean of y so that $\bar{y} = \sum v_i / g$ estimates it consistently. τ_i^2 is estimated consistently by t_1^2

as seen from (113) of Chapter III. Let $\hat{\rho}\tau$ be some consistent estimate of $\rho\tau$. Then $\hat{\Lambda}$ can be estimated consistently by the use of (4), (5), (28) and (29).

5.4. Model sampling for limit distributions in fixed interval analysis.

In sections 4.5, and the previous sections we made use of the several limit theorems of section 4.4. The rate at which these limit distributions are reached needs to be studied. With this view we have conducted some model sampling experiments, with samples sizes of the order of 50, to compare the sampling distribution with the limit distribution. We present these results below.

The first experiment is one connected with fixed interval analysis with $g = 2$; $a_0 = -\infty$, $a_1 = 0$, $a_2 = +\infty$. To simplify matters we have taken $Y = X$. In this case Y is completely dependent on X , and the limiting distributions are perhaps attained at a slower rate than when Y and X are less dependent. The distribution of X is that of a standard normal variable. 208 samples each with $n = 50$ were taken from the table for normal random deviates by Sengupta and Bhattacharya (1958). \bar{v}_1 and \bar{v}_2 were calculated from each of these samples. We see that, here

$$-\bar{v}_1 = \bar{v}_2 = \sqrt{\frac{2}{\pi}}, \quad \bar{v}_1^2 = \bar{v}_2^2 = 1 \quad \text{and} \quad \kappa_1 = \kappa_2 = \frac{1}{2}. \quad \text{Thus if}$$

$$\bar{\eta}_1 = \sqrt{25} (\bar{v}_1 + \sqrt{2/\pi}) \quad \dots \quad (31)$$

$$\bar{\eta}_2 = \sqrt{25} (\bar{v}_2 - \sqrt{2/\pi}) \quad \dots \quad (32)$$

then according to theorem 4.4.5, the distribution of $(\bar{\eta}_1, \bar{\eta}_2)$

converges to the distribution of two independent standard normal variables. The frequency distribution of this statistic as observed from the 208 samples is given in table V as a two way table together with the expected frequencies. The observed value of the discrepancy χ^2 with 24 d.f. is 26.615. This shows that there is a good conformity with the limiting distribution. We feel it plausible therefore, that the sample size required for the limiting distributions to be reached in fixed interval analysis is of the same order as that required for the approximation using the central limit theorem.

5.5. Model sampling experiments in fractile analysis.

The vector (Y, X) here has a bivariate normal distribution with parameters $0, 0; 2, 1; \frac{1}{\sqrt{2}}$. Samples S and S' with $n = 48$, $n' = 48$ were drawn independently. In the first instance g was taken to be 4 so that $m = 12$. v_1, v_2, v_3 and v_4 were computed from S and v'_1, v'_2, v'_3 and v'_4 from S' . This experiment was repeated 148 times. These bivariate random observations were taken by making use of the tables of random normal deviates of Sengupta and Bhattacharya (1958), Wold (1958) and RAND corporation (1955). The following statistics were computed

$$\Delta^0 = \frac{m}{2d} \sum (v_i - v'_i)^2 \quad \dots \quad (33)$$

where $2d = 3.035808$

$$\Gamma = m (\underline{v} - \underline{v}') N^{-1} (\underline{v} - \underline{v}')' \quad \dots \quad (34)$$

where $N = 4 I + 2 Q_4$, Q_4 being the matrix Q obtained from table I with $g = 4$. The frequency distributions of these statistics are given in tables VI and VII. According to theory the limiting distribution of Δ^0 is that of

$$1.517606 \chi_1^2 + 0.703727 \chi_4^2 + 0.955632 \chi_3^2 + 0.787400 \chi_2^2$$

where $\chi_1^2, \dots, \chi_4^2$ are all independent chisquares with 1 d.f.

This is got by using the latent roots of N which can be computed with the use of table II. This distribution can be approximated by a χ^2 with 3.763912 d.f. as can be seen by using (8) and (9). The limiting distribution of Γ is a χ^2 with 4 d.f. From the values of the discrepancy χ^2 in tables VI and VII we conclude that these sampling and limiting distributions are in close agreement.

The experiment was repeated with the same samples by taking $g = 2$, $m = 24$. The two statistics that were computed are

$$\Delta^0 = \frac{m}{2d} ((v_1 - v_1')^2 + (v_2 - v_2')^2) \quad \dots \quad (35)$$

where $2d = 2.726762$.

$$\Gamma = m (\underline{v} - \underline{v}') N^{-1} (\underline{v} - \underline{v}')' \quad \dots \quad (36)$$

where $N = 4I + 2Q_2$, where Q_2 is the Q matrix obtained from table I with $g = 2$. The frequency distributions of these statistics are given in tables VIII and IX. The limiting distribution of Δ^0 is that of $\chi_1^2 + 1.466941 \chi_2^2$ where χ_1^2 and χ_2^2 are independent χ^2 with 1 d.f. The limiting distribution function has been approximated by a sum of distribution functions of chisquares, using an approximation due to Robbins and Pitman (1949). The limiting distribution of Γ is a χ^2 with 2 d.f. We find from tables VIII and IX that, once again, the empirical and limiting distributions are in agreement.

Table I.

The following are the Q matrices Q_g for $g = 2(1)16, 20, 25$. $(Q_g)_{ij}$ has been given only for $j \leq i \leq g-j+1$, $1 \leq j \leq [\frac{g+1}{2}]$, since Q is doubly symmetric. (See (20).) The elements of the matrices Q_g for $g = 16, 20$ and 25 may be in error in the last decimal place. These errors have been carried on to tables II and III.

Table I.

		$10^5 \times (Q_g)_{ij}$		
	j	1	2	3
	i	(g = 2)		
1		-31831		
2		31831		
<hr/>				
	i	(g = 3)		
1		-42954		
2		28431	-56862	
3		14523		
<hr/>				
	i	(g = 4)		
1		-49140		
2		25339	-65199	
3		14903	24958	
4		8899		
<hr/>				
	i	(g = 5)		
1		-53210		
2		23160	-69826	
3		14141	21961	-72205
4		9677	15028	
5		6231		

Table I (contd.)

 $10^5 \times (Q_g)_{ij}$

j	1	2	3	4
1	(g = 6)			
1	-56149			
2	21549	-72858		
3	13389	19854	-76069	
4	9513	14107	19207	
5	6987	10361		
6	4712			
<hr/>				
i	(g = 7)			
1	-58399			
2	20301	-75034		
3	12735	18296	-78601	
4	9216	13240	17285	-79482
5	7023	10089	13172	
6	5379	7728		
7	3744			
<hr/>				

Table I (contd.)

$$10^5 \times (Q_g)_{ij}$$

i	j	1 ($g = 8$)	2	3	4	5
1		-60194				
2		19302	-76689			
3		12179	17091	-80419		
4		8907	12500	15866	-81711	
5		6913	9701	12314	15509	
6		5489	7703	9777		
7		4324	6068			
8		3081				
<hr/>						
i		($g = 9$)				
1		-61670				
2		18477	-78001			
3		11702	16127	-81804		
4		8617	11875	14769	-83312	
5		6761	9317	11588	14202	-83736
6		5466	7532	9368	11482	
7		4460	6146	7644		
8		3586	4941			
9		2602				

Table I (contd.)

$$10^5 \times (q_g)_{ij}$$

j	1	2	3	4	5
	(g = 10)				
	-62913				
	17783	-79073			
	11290	15333	-82902		
	8352	11343	13891	-84531	
	6600	8963	10976	13193	-85187
	5394	7326	8971	10782	12982
	4480	6084	7451	8955	
	3727	5062	6199		
	3044	4134			
	2241				

Table I (contd.)

$$10^4 \times (Q_g)_{ij}$$

j	1	2	3	4	5	6
1	(g = 11)					
1	-6398					
2	1719	-7997				
3	1093	1467	-8380			
4	811	1089	1317	-8550		
5	644	864	1046	1238	-8629	
6	530	712	861	1019	1204	-8653
7	445	597	723	856	1011	
8	377	506	612	725		
9	318	427	517			
10	263	353				
11	196					

Table I (contd.)

$$10^4 \times (a_g)_{ij}$$

j	1	2	3	4	5	6
(g = 12)						
	-6491					
	1667	-8073				
	1061	1410	-8455			
	790	1049	1256	-8629		
	629	836	1001	1172	-8717	
	521	692	828	970	1129	-8755
	440	585	700	820	955	1115
	377	500	599	702	817	
	324	430	515	603		
	277	367	440			
	230	307				
	174					

Table I (contd.)

 $10^4 \times (Q_g)_{ij}$

j	1	2	3	4	5	6	7
(g = 13)							
	-6573						
	1622	-8141					
	1033	1361	-8518				
	770	1014	1204	-8695			
	615	810	962	1116	-8788		
	511	673	799	927	1067	-8836	
	434	572	679	788	907	1045	-8850
	374	493	585	679	782	900	
	325	428	508	589	678		
	282	372	441	512			
	243	321	381				
	205	270					
	156						

Table I (contd.)

$$10^4 \times (q_g)_{ij}$$

1	2	3	4	5	6	7
(g = 14)						
-6645						
1581	-8197					
1008	1317	-8574				
752	983	1159	-8751			
603	787	928	1068	-8848		
502	656	773	890	1015	-8902	
428	559	660	759	866	986	-8926
371	485	572	656	750	855	977
324	424	500	575	656	747	
285	372	439	505	576		
250	326	385	443			
217	283	334				
184	240					
141						

Table I (contd.)

$$10^4 \times (Q_g)_{ij}$$

j	1	2	3	4	5	6	7	8
(g = 15)								
	-6711							
	1545	-8248						
	985	1279	-8622					
	736	955	1120	-8800				
	590	766	898	1027	-8899			
	493	640	750	857	970	-8957		
	422	547	642	733	830	938	-8987	
	367	476	558	638	722	815	922	-8997
	322	418	590	560	635	716	810	
	285	370	433	495	561	633		
	252	327	384	439	497			
	223	289	339	388				
	195	253	297					
	166	216						
	128							

Table I (contd.)

$$10^4 \times (Q_g)_{1j}$$

1	2	3	4	5	6	7	8
(g = 16)							
-6770							
1513	-8294						
965	1244	-8665					
721	903	1085	-8842				
579	747	871	990	-8943			
484	625	729	828	932	-9004		
415	536	625	710	799	896	-9039	
362	468	545	620	697	781	876	-9055
320	412	481	547	615	689	773	870
284	366	427	486	546	612	687	
253	327	381	433	487	546		
226	292	340	386	435			
201	259	302	344				
177	228	266					
152	196						
118							

Table II

Latent roots $(q_{g,1}, \dots, q_{g,g}) = q_g$ of the Q matrices: $g=2(1)16,20,25$.

$$10^5 \times q_g$$

$$g= 2$$

$$(00000, -63662)$$

$$g= 3$$

$$(00000, -85294, -57478)$$

$$g= 4$$

$$(00000, -93249, -80483, -54946)$$

$$g= 5$$

$$(00000, -96323, -77658, -90674, -53621)$$

$$g= 6$$

$$(00000, -97716, -52822, -94976, -88817, -75820)$$

$$g= 7$$

$$(00000, -96960, -52293, -98448, -93896, -87424, -74527)$$

$$g= 8$$

$$(00000, -51920, -98876, -97989, -86339, -96315, -93020, -73569)$$

$$g= 9$$

$$(00000, -51644, -99148, -85466, -98579, -97583, -95767, -92295, -72828)$$

$$g=10$$

$$(00000, -51433, -98945, -84746, -99332, -98309, -97226, -95297, -91684, -72239)$$

Table II (contd.)

$$10^4 \times q_g$$

g-11

(0000, -5127, -9876, -8414, -9946, -9116, -9919, -9807, -9691, -9489,
-7176)

g-12

(0000, -5113, -9859, -8362, -9956, -9070, -9935, -9905, -9785, -9664,
-9452, -7136)

g-13

(0000, -5102, -9843, -8317, -9963, -9030, -9948, -9926, -9422, -9892,
-9766, -9639, -7102)

g-14

(0000, -5093, -9881, -8278, -9968, -8994, -9937, -9393, -9940, -9916,
-9829, -9748, -9616, -7073)

g-15

(0000, -5085, -9933, -8243, -9972, -8962, -9950, -9596, -9964, -9871,
-9908, -9816, -9732, -9367, -7048)

g-16

(-0003, -5082, -9901, -8213, -8936, -9976, -9969, -9570, -9946, -9710,
-9959, -9928, -9861, -9801, -9340, -7025)

g-20

(0001, -5060, -9905, -6960, -8836, -9972, -9954, -9653, -10003, -9267,
-9983, -9735, -9977, -9962, -9945, -9926, -9874, -9823, -9541, -8116)

Table II (contd.)

$$10^4 \times q_g$$

g-25

(0003, -5040, -9940, -6908, -8757, -9991, -9456, -9968, -9799, -9989,
-9925, -9987, -9973, -9727, -9849, -9986, -9976, -9960, -9952, -9882,
-9981, -9906, -9621, -9192, -8038)

Table III

Trace $\text{Tr } Q$, of the Q matrices; $g = 2(1)16, 20, 25$.

g	$-\text{Tr } Q_g$
2	0.63662
3	1.42770
4	2.28678
5	3.18277
6	4.10152
7	5.30550
8	5.98026
9	6.93310
10	7.82212
11	8.8561
12	9.8240
13	10.7952
14	11.7686
15	12.7485
16	13.7224
20	17.6794
25	22.5800

Table IV.

The following gives $\sum_1^g \mu_{g,i}^2$ for $g=2(1)16,20,25$.

g	$\sum_1^g \mu_{g,i}^2$
2	0.636620
3	2.379689
4	3.442239
5	4.484782
6	5.516166
7	6.540582
8	7.560287
9	8.576615
10	9.590464
11	10.602399
12	11.612864
13	12.622234
14	13.630309
15	14.637715
16	15.644487
20	19.666152
25	24.685734

Table V

Model sampling experiments in fixed interval analysis.

$Y = X$, X is normal $(0, 1)$; $n = 50$; $g = 2$; $a_0 = -\infty$, $a_1 = 0$, $a_2 = +\infty$.

Statistics $(\bar{\eta}_1, \bar{\eta}_2)$ are defined in (31), (32). See section 4.

In the following table, [1] represents observed frequency and [2] represents the frequency of the limiting distribution.

$\bar{\eta}_2$ \ / \ $\bar{\eta}_1$		$(-\infty)$ to -0.8416	-0.8416 to -0.2534	-0.2534 to -0.2534	0.2534 to 0.8416	0.8416 to ∞
$-\infty$ to -0.8416	[1]	7	14	5	10	6
-0.8416 to -0.8416	[2]	8.32	8.32	8.32	8.32	8.32
-0.8416 to -0.2534	[1]	11	6	7	6	13
-0.2534 to -0.2534	[2]	8.32	8.32	8.32	8.32	8.32
-0.2534 to 0.2534	[1]	4	4	6	11	8
0.2534 to 0.2534	[2]	8.32	8.32	8.32	8.32	8.32
0.2534 to 0.8416	[1]	8	11	6	9	3
0.8416 to 0.8416	[2]	8.32	8.32	8.32	8.32	8.32
0.8416 to ∞	[1]	9	12	12	9	11
∞ to ∞	[2]	8.32	8.32	8.32	8.32	8.32

Total = 208

Calculated discrepancy $\chi^2 = 26.615$; d.f. = 24

5% critical discrepancy = 36.415

Table VI

Model sampling experiments in fractile analysis.

(Y, X) is bivariate normal $(0,0; 2, 1; 1/\sqrt{2})$; $n=48$, $g=4$, $m=12$.

Statistic $z = \Delta^\circ$ defined in (33). See section 5.

z	observed frequency	frequency of the approximate limiting distribution *.
0.0 - 0.7	4	9.472
- 1.1	13	9.620
- 1.4	5	8.436
- 1.7	9	8.436
- 1.9	7	5.772
- 2.2	7	7.844
- 2.4	6	5.920
- 2.8	14	10.212
- 3.0	7	5.772
- 3.4	8	8.880
- 3.6	5	4.588
- 4.0	12	8.140
- 4.4	4	7.252
- 4.8	5	6.808
- 5.4	10	7.844
- 6.0	5	7.104
- 6.8	4	6.956
- 7.8	9	6.216
- 9.4	8	6.364
- ∞	6	6.364
Total	148	148.000

Calculated discrepancy $\chi^2 = 15.858$; d.f. = 19.

5% critical discrepancy = 30.1435.

* i.e. a χ^2 distribution with 3.763912 d.f.

Table VII

Model sampling experiments in fractile analysis.

(Y, X) is bivariate normal $(0,0; 2,1, 1/\sqrt{2})$; $n=48$, $g=4$, $m=12$.

Statistic $z = \Gamma$ is defined in (34). See section 5.

z	observed frequency	frequency of the limiting distribution
0.0 - 0.7	4	7.203
- 1.1	9	8.445
- 1.4	8	7.410
- 1.7	6	7.915
- 1.9	5	5.414
- 2.2	8	8.156
- 2.4	5	5.387
- 2.8	11	10.478
- 3.0	8	5.032
- 3.4	11	9.558
- 3.6	3	4.501
- 4.0	14	8.411
- 4.4	4	7.613
- 4.8	5	6.827
- 5.4	11	8.847
- 6.0	5	7.327
- 6.8	8	7.742
- 7.8	4	7.052
- 9.4	11	7.008
- ∞	8	7.672
Total	148	147.998

Calculated discrepancy $\chi^2 = 15.330$; d.f. = 19

5% critical discrepancy 30.1435.

Table VIII

Model sampling experiments in fractile analysis.
 (Y,X) is bivariate normal $(0,0; 2,1; 1/\sqrt{2})$; $n=48$, $g=2$, $m=24$.
 Statistic $z = \Delta^{\circ}$ is defined in (35). See section 5.

z	observed frequency	frequency of the limiting distribution *
0.0 - 0.1	5	5.983
- 0.2	4	5.737
- 0.3	2	5.501
- 0.4	7	5.271
- 0.6	12	10.210
- 0.7	4	4.356
- 0.9	10	8.741
- 1.0	3	4.104
- 1.2	7	7.712
- 1.4	6	7.095
- 1.6	3	6.530
- 1.8	4	6.009
- 2.0	9	5.273
- 2.4	13	10.035
- 2.8	9	8.288
- 3.2	8	7.036
- 3.6	10	8.597
- 4.6	9	8.634
- 6.0	6	9.759
- ∞	17	13.128
Total	148	147.999

Calculated discrepancy $\chi^2 = 14.531$; d.f. = 19.

5% critical discrepancy = 30.1435.

* The limiting distribution function can be written as

$$\begin{aligned}
 & 0.82564550 F_2(x) + 0.13140540 F_4(x) + 0.03137069 F_6(x) + 0.00832132 F_8(x) \\
 & + 0.00832132 F_8(x) + 0.00231766 F_{10}(x) + 0.00066395 F_{12}(x) \\
 & + 0.00019373 F_{14}(x) + 0.00005726 F_{16}(x) + 0.00001708 F_{18}(x) \\
 & + 0.00000513 F_{20}(x) + R(x)
 \end{aligned}$$

where $0 \leq R(x) \leq 0.00000227$ for all x and $F_p(x)$ is the distribution function of a chi-square distribution with

Table IX

Model sampling experiments in fractile analysis.

(Y,X) is bivariate normal $(0,0; 2,1; 1/\sqrt{2})$; $n=48, g=2, m=24$.

Statistic $z = \Gamma$ is defined in (35). See section 5.

z	observed frequency	frequency of the limiting distribution
0.0 - 0.1	6	7.218
- 0.2	3	6.866
- 0.3	6	6.531
- 0.4	8	6.213
- 0.6	14	11.531
- 0.7	8	5.347
- 0.9	8	9.925
- 1.0	6	4.603
- 1.2	5	8.543
- 1.4	4	7.729
- 1.6	7	6.994
- 1.8	6	6.328
- 2.0	8	5.726
- 2.4	9	9.870
- 2.8	16	8.079
- 3.2	8	6.616
- 3.8	6	7.745
- 4.6	3	7.298
- 6.0	7	7.470
- ∞	10	7.369
Total	148	148.001

Calculated discrepancy $\chi^2 = 21.795$; d.f. = 19.

5% critical discrepancy 30.1435.

LIST OF REFERENCES

- [1] Billingsley, P. (1956): The invariance principle for dependent random variables, Trans. Amr. Math. Soc., 83, 250-268.
- [2] Box, G.E.P. (1949): A general distribution theory for a class of likelihood criteria, Biometrika, 36, 317-346.
- [3] Cramer, H. (1937): Random variables and probability distributions, Cambridge University Press.
- [4] Cramer, H. (1946): Mathematical methods of statistics, Princeton University Press.
- [5] Das, Rhea S. (1960): Applications of fractile graphical analysis to psychometry: I. Item analysis, Psychological studies, 5, 11-18.
- [6] Das, Rhea S. and Sharma, K. N. (1960): Applications to fractile graphical analysis to psychometry: II. Reliability, Psychological Studies, 5, 71-77.
- [7] Davis, H. T. (1941): The theory of econometrics, The Principia Press, Bloomington, Indiana.
- [8] Doob, J. L. (1953): Stochastic processes, John Wiley and sons, New York.
- [9] Dunford, N. and Schwartz, J. T. (1958): Linear operators, Part I, Interscience publishers, New York.
- [10] Fisher, R. A. (1922): On the interpretation of chisquare from contingency tables and in the calculation of P, Jour. Royal. Stat. Soc., 85, 87-94.
- [11] Fisher, R. A. and Tippett, L.H.C. (1928): Limiting forms of the frequency distribution of the largest or smallest number of a sample, Proceed. Camb. Phil. Sec., 24, 180-190.
- [12] Fraser, D.A.S. (1957): Non-parametric methods in statistics, John Wiley and sons, New York.
- [13] Fréchet, M. (1927): Sur la loi de probabilité de l'écart maximum, Annales de la société polonaise de Mathématiques, 6, 93.

- [14] Gihman, I. I. (1953): Some limit theorems for conditional distributions and on problems of mathematical statistics connected with them, Ukrain Math. Zhurnal, 5, 413-433 (Russian).
- [15] Gnedenko, B. V. (1943): Sur la distribution limite du terme maximum d'une série aléatoire, Ann. Math. 44, 423-453.
- [16] Gnedenko, B. V. and Kolmogorov, A.N. (1954): Limit distributions for sums of independent random variables: Translated from the Russian by K.L.Chung, Addison-Wesley Publishers, Cambridge.
- [17] Gosh, B.N. (1949): Interpenetrating (networks of) samples, Cal. Stat. Assoc. Bull. 2, 108-119.
- [18] Gumbel, E.J. (1958): Statistics of extremes, Columbia University Press.
- [19] Halmos, P.R. (1950): Measure theory, Van Nostrand Company, New York.
- [20] Heeffding, W. (1948): On a class of statistics with asymptotically normal distributions, Ann. Math. Stat. 19, 293-325.
- [21] Heeffding, W. (1953): On the distribution of the expected values of the order statistics, Ann. Math. Stat., 24, 93-100.
- [22] Iyengar, N. S. (1960): On a method of computing Engel elasticities from concentration curves, Econometrica, 28, 862-891.
- [23] Kawata, T. (1951): Limit distributions of single order statistics, Rep. Stat. Appl. Res. Union of Japanese scientists and engineers, 1, 4-9.
- [24] Kendall, M.G. and Stuart, A. (1958): The advanced theory of statistics, I, Charles Griffin and Co., London.
- [25] Lahiri, D.B. (1954): Technical paper on some aspects of the development of the sample design, National Sample Survey of India, Report No.5, Indian Statistical Institute, Calcutta.
- [26] Lahiri, D.B. (1957): Observations on the use of interpenetrating samples in India, Bulletin de l'institut international de statistique, 36, Book 3, 144-152.
- [27] Lehmann, E. (1951): Consistency and unbiasedness of certain non-parametric tests, Ann. Math. Stat., 22, 165-179.

- [28] Lévy, P. (1937): Théorie de l'addition des variables aléatoires, Monographies des probabilités, Fasc. 1, Paris.
- [29] Mahalanobis, P.C. (1946): Recent experiments in statistical sampling in the Indian Statistical Institute, Jour. Royal Stat. Soc. 109, 325-378.
- [30] Mahalanobis, P.C. (1958a): Lectures in Japan - draft proof.
- [31] Mahalanobis, P.C. (1958b): A method of fractile analysis with some surmises of results, Transactions of the Bose Institute, 22, 223-230.
- [32] Mahalanobis, P.C. (1960): A method of fractile graphical analysis, Econometrica, 28, 325-351. (Reprinted in Sankhya Series A, 23, (1961), 41-64.
- [33] V. Mises, R. (1947): On the asymptotic distribution of differentiable statistical functions, Ann. Math. Stat. 18, 309-348.
- [34] Munroe, M.E. (1953): Introduction to measure and integration, Addison Wesley Publishing company, Cambridge, Mass.
- [35] Parthasarathy, K.R. and Bhattacharya, P.K. (1961): Some limit theorems in regression theory, Sankhya, Series A, 23, 91-102.
- [36] Pargen, E. (1954): On uniform convergence of families of sequences of random variables, University of California Publications in Statistics, 2, No.2, 23-54.
- [37] Plackett, R.L. (1960): Principles of Regression Analysis, Oxford University Press.
- [38] Prohorov, Yu. V. (1956): Convergence of random processes and limit theorems in the theory of probability, Teoreya Veroyat. i. Primen. 1, 177-238 (Russian).
- [39] RAND Corporation, The (1955): A million random digits with 100,000 normal deviates, Free Press, Illinois.
- [40] Ranga Rao, R. (1960): Some problems in probability theory, Unpublished D.Phil. Thesis, submitted to the University of Calcutta.

- [41] Ranga Rao, R. (1961): Some theorems on weak convergence of measures and its applications, Submitted to Ann. Math. Stat.
- [42] Robbins, H. and Pitman, E.J.G. (1949): Application of a method of mixtures to quadratic forms in normal variates, Ann. Math. Stat., 20, 552-560.
- [43] Roy, J., Chakravarty, I.M. and Laha, R.G. (1959): A study of concentration curves as description of consumption pattern, Studies relating to planning for national development, 2, 73-82. Indian Statistical Institute, Calcutta.
- [44] Roy, J. and Dhar, S.K. (1959): A study of the pattern of consumer expenditure in rural and urban India, Studies relating to planning for national development 3, 56-72, Indian Statistical Institute, Calcutta.
- [45] Satterthwaite (1946): An approximate distribution of estimates of variance components, Biometrics, 2, 110-114.
- [46] Scheffé, H. (1947): A useful convergence theorem for probability distributions, Ann. Math. Stat., 18, 434-438.
- [47] Sengupta, J.M. and Bhattacharya, N. (1958): Tables of Random Normal Deviates, Sankhya, 20, 250-286.
- [48] Sethuraman, J. (1961): Some limit theorems connected with fractile graphical analysis, Sankhya, Series A, 23, 79-90.
- [49] Sethuraman, J. and Sukhatme, ^{e.v.} (1959): Joint asymptotic distribution of U-statistics and order statistics, Sankhya, 21 289-298.
- [50] Smirnov, N.V. (1949): Limit distributions for the terms of a variational series, Trudy. Mat. Inst. A. Steklova, 25, 1-61. (Russian).
- [51] Smirnov, N.V. (1952): Limit distributions for the terms of a variational series, Ann. Math. Soc. Translation No.67.
- [52] Som, R.K. (1960): A statistical study of population in India, Unpublished D.Phil. Thesis, submitted to the University of Calcutta.

- [53] Steck, G.P. (1957): Limit theorems for conditional distributions, University of California publications in statistics, No.12, 237-284.
- [54] Sukhatme, B.V. (1957): Joint asymptotic distribution of the median and a U-statistic, Jour. Royal Stat. Soc., Series B, 19, 144-148.
- [55] Sukhatme, B.V. (1958): Testing the hypothesis that two populations differ only in location, Ann. Math. Stat. 29, 60-78.
- [56] Takeuchi, K. (1961): On some properties of error area in the fractile graph method, Sankhya, Series A, 23, 65-78.
- [57] Varadarajan, V.S. (1958): A useful convergence theorem, Sankhya, 20, 221-222.
- [58] Varadarajan, V.S. (1958a): Convergence of stochastic processes, unpublished D. Phil., Thesis, submitted to the University of Calcutta.
- [59] Williams, E.J. (1959): Regression Analysis, John Wiley and sons, New York.
- [60] Weld, H.O.A. (1948): Random Normal Deviates, Tracts for computers No.XXV, Cambridge University Press.
- [61] Yates, F. (1949): Sampling methods for censuses and surveys, Charles Griffin and Company, London.

