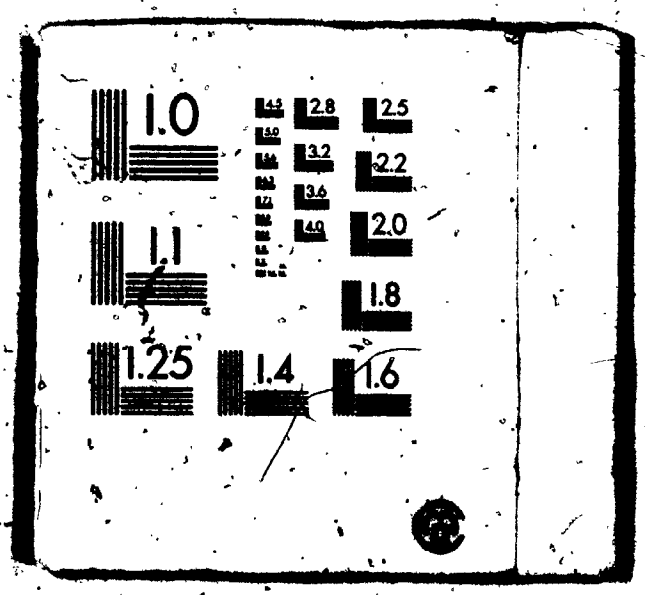
1



52746

0-315-04852-2

National Library of Canada | Bibliothèque nationale du Canada | CANADIAN THESES ON MICROFICHE | THÈSES CANADIENNES SUR MICROFICHE

NAME OF AUTHOR/NOM DE L'AUTEUR EMAD-ELDIN ALY AHMED ALY

TITLE OF THESIS/TITRE DE LA THÈSE QUANTILE PROCESS, SPACINGS AND GOODNESS-OF-FIT

UNIVERSITY/UNIVERSITÉ CARLETON UNIV.

DEGREE FOR WHICH THESIS WAS PRESENTED/ GRADE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE Ph.D.

YEAR THIS DEGREE CONFERRED/ANNÉE D'OBTENTION DE CE DEGRÉ 1981

NAME OF SUPERVISOR/NOM DU DIRECTEUR DE THÈSE M. CSÖRGŐ

DATED/DATE May 7, 1981 SIGNED/SIGNÉ Emad Eldin Aly Ahmed Aly

PERMANENT ADDRESS/RÉSIDENCE FIXE (up to sept. 1983) Dept. of Math., Carleton Univ., Ottawa, K1S 5B6

(After sept. 1983) 9B Saeed street, ELSKAKINY ABBASSIA, CAIRO, EGYPT.

NL-91 (11-73)

QUANTILE PROCESS, SPACINGS AND GOODNESS-OF-FIT

by

Emad-Eldin Aly Ahmed Aly, B.Sc. (Hons.),

M.Sc., M.Sc.

A thesis submitted to the Faculty of
Graduate Studies and Research in Partial
fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Mathematics and Statistics

Carleton University

Ottawa, Canada

April 1981

The undersigned recommend to the Faculty of Graduate Studies and Research acceptance of the thesis

"Quantile Process, Spacings and Goodness-of-Fit",

submitted by Emad-Eldin Aly Ahmed Aly, B.Sc. (Hons.), M.Sc., M.Sc., in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Miklós Csörgő

.....

Dr. M. Csörgő, Thesis Supervisor

K. S. Williams

.....

Dr. K.S. Williams, Chairman, Department of Mathematics and Statistics

.....

Dr. Constance Van Eeden, External Examiner

Carleton University

April 1981

ABSTRACT

Based on the strong approximation results for the quantile process of Csörgő, Révész (1978), the so-called "quantile approach" is studied. This approach is useful for goodness-of-fit statistical tests in the presence of location and scale parameters, i.e., when testing for $H_o : F(x;\mu,\sigma) = F_o(\frac{x-\mu}{\sigma})$, with F_o specified and $-\infty < \mu < +\infty$, $\sigma > 0$ are nuisance location and scale parameters. Estimating (μ, σ) by a "reasonable" sequence of estimators $(\hat{\mu}_n, \hat{\sigma}_n)$, we derive the aymptotic distribution of the following Cramér-von Mises type statistics

$$M_n(\lambda) = \sum_{i=1}^{n} \left\{ \frac{X_{i:n} - \hat{\mu}_n}{\hat{\sigma}_n} - F_o^{-1}(\frac{i}{n+1}) \right\}^2 f_o(F_o^{-1}(\frac{i}{n+1}))(F_o^{-1}(\frac{i}{n+1}))^{\lambda-1},$$

where $\lambda \geq 1$ is a fixed integer and $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ are the order statistics of the given random sample. In addition, we also derive an asymptotic in-probability representation of the estimated standardized quantile process, $\hat{\beta}_n(.)$, when estimating (μ, σ) by a "reasonable" sequence of estimators $(\hat{\mu}_n, \hat{\sigma}_n)$.

The strong approximation results of Komlós, Major, Tusnády are used to derive the asymptotic distribution of some stochastic processes defined in terms of uniform and exponential spacings and

the corresponding rates of convergence are also given. In addition, our general treatment enables us to introduce and discuss some new nuisance parameter free distribution free goodness-of-fit tests for the uniform and exponential distributions.

The asymptotic distribution of a test for clustering is also obtained via the "quantile approach". These results are generalization of the results of Brillinger, Knott, Scott (1979).

Throughout the dissertation emphasis is placed not only on the particular results obtained, but also on the strong approximation methodology used.

To the memory of

three I loved most,

my parents and my

brother Naser.

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation for the help, guidance and constant encouragement of my thesis supervisor, Professor Miklós Csörgő, who enriched my knowledge in Mathematics and shared his ideas unselfishly with me.

Mrs. F. Zaydan, who typed this thesis, deserves a great deal of credit for her fast and accurate work.

The financial assistance provided by Carleton University in the form of Scholarships and Teaching/Research Assistantships is gratefully appreciated.

Lastly, but not least, I wish to express my appreciation to my wife, Faten, who was a constant and understanding support during the preparation of this thesis.

TABLE OF CONTENTS

FOREWORD

A classic problem in statistics is that of goodness-of-fit. Based on the empirical process of a random sample of size n, the Kolmogorov-Smirnov and the Cramér-von Mises functionals can be used to test the simple goodness-of-fit hypothesis that our random sample has a completely specified distribution function. However, in most cases, one only knows the form of the distribution function, $F(x;\theta)$, while some parameters θ are not specified. There are many ways of "getting rid of θ" so as to reduce composite goodness-of-fit hypotheses to simple ones. As far as the empirical process is concerned, one natural way of doing this is to estimate θ by using some kind of a "good estimator" sequence $\{\theta_n\}_{n=1}^{\infty}$ based on the given random sample from $F(x;\theta)$. For illumination of these remarks as well as further results and references, we refer to Burke et al (1979), Darling (1955), Durbin (1973a), Kac, Kiefer, Wolfowitz (1955) and Neuhaus (1974, 1976).

A common drawback of the general results of the above quoted papers is that the resulting Gaussian processes which approximate the estimated empirical process depend, in general, on both the form of the hypothetical distribution function and the true values of the unknown parameters of the latter.

One way to overcome the above mentioned drawback is the so-called random substitution method suggested by Durbin (1961, 1976). Another randomization approach is the so-called half-sample device suggested by Durbin (1973, 1976) (also cf. Burke et al (1979)). The main drawback of these randomization approaches is that it introduces into the analysis of real data the element of artificial randomization and it is suspected that a loss of power may result via the use of randomization.

A completely different, but related, approach is the so-called "quantile approach" based on the quantile process of a given random sample of size n. This approach is based on the strong approximation results of Csörgő, Révész (1978) for the quantile process. Again the Kolmogorov-Smirnov and the Cramér-von Mises functionals of the quantile process can be used to test the simple goodness-of-fit hypothesis that our random sample has a completely specified D.F.

Concerning the location and scale parameters family of D.F.'s, the "quantile approach" turned out to be very useful in dealing with the above composite goodness-of-fit tests. Csörgő, Révész (1979, 1981) considered the composite goodness-of-fit hypothesis $H_o : F(x,\mu,\sigma) = F_o(x - \mu)/\sigma)$ with F_o specified and $-\infty < \mu < +\infty$, $\sigma > 0$ are unknown location and scale parameters. They proposed the following test statistic:

$$M_n(\lambda) = \sum_{k=1}^{n} f_o(F_o^{-1}(\frac{k}{n+1}))\left\{\frac{X_{k:n} - \hat{\mu}_n}{\hat{\sigma}_n} - F_o^{-1}(\frac{k}{n+1})\right\}^2 (F_o^{-1}(\frac{k}{n+1}))^{\lambda-1},$$

where $\lambda \geq 1$ is a fixed integer, $f_o = F_o'$, F_o^{-1} is the inverse of F_o, $(\hat{\mu}_n \ \hat{\sigma}_n)$ is a "reasonable" sequence of estimators of $(\mu \ \sigma)$ and $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ are the order statistics of the given random sample $X_1, X_2, \ldots, X_n$. Using the "quantile approach" and if $(\hat{\mu}_n \ \hat{\sigma}_n) = (\bar{X}_n \ S_n)$, where $\bar{X}_n$ is the sample mean and S_n^2 is the sample variance, Csörgő, Révész (1979, 1981) derived the asymptotic distribution of $M_n(\lambda)$, which ~~turned out to be~~ nuisance parameter free but it depends, in general, on the form of F_o.

As an alternate route for testing the same composite goodness-of-fit hypothesis, Csörgő, Révész (1980 I,II), following up a conjecture by Parzen (1979a,b), suggested tests based on

$$p_n(u) = \begin{cases} 0 & \text{if } o \leq u \leq \frac{2}{n} \\ n^{1/2}\left(\dfrac{\sum_{1}^{[nu]-1} f_o(F_o^{-1}(\frac{k}{n+1}))(X_{k+1:n} - X_{k:n})}{\sum_{1}^{n-1} f_o(F_o^{-1}(\frac{k}{n+1}))(X_{k+1:n} - X_{k:n})} - u\right) & \text{if } \frac{2}{n} \leq u \leq 1, \end{cases}$$

where $[x]$ stands for the integer part of the real number x. Again, via the "quantile approach" Csörgő, Révész (1980 I, II) derived the asymptotic distribution of $p_n(u)$ which turned out to

be nuisance parameter free but it depends on the form of F_o, except in the uniform and exponential cases.

Another alternative route to the same problem is to consider tests based on

$$g_n(u;\alpha) = \begin{cases} \text{if } 0 \le u < \frac{2}{n} \\ n^{1/2}\left[\dfrac{n^{\alpha-1}\sum_{1}^{[nu]-1} f_o^{\alpha}(F_o^{-1}(\frac{k}{n+1}))(X_{k+1:n} - X_{k:n})^{\alpha}}{\left(\sum_{1}^{n-1} f_o(F_o^{-1}(\frac{k}{n+1}))(X_{k+1:n} - X_{k:n})\right)^{\alpha}} - u\Gamma(1+\alpha)\right] \\ \text{if } \frac{2}{n} \le u \le 1 , \end{cases}$$

where $\alpha > 0$. It appears that deriving the asymptotic distribution of $g_n(u;\alpha)$ via the "quantile approach" is extremely difficult in general for $\alpha \neq 1$, including the important special case of $\alpha = 2$. However, another strong approximation methodology, namely, that of Komlós, Major, Tusnády (1975, 1976), turned out to be very useful in the uniform and exponential cases for any $\alpha > 0$. The problem of deriving the asymptotic distribution of $g_n(u;\alpha)$ for distributions other than the uniform and exponential ones for $\alpha \neq 1$ remains open.

Another classic problem that can be treated via the "quantile approach" is that of clustering techniques, partitioning the observations into two groups so that the between groups sum of squares is maximized.

This dissertation is divided into five chapters. In chapter I, we derive in-probability representation of an alternative definition of the sample quantile function. This definition is suggested by Parzen (1979a,b).

Chapter II deals with obtaining the asymptotic distribution of $M_n(\lambda)$ via the "quantile approach", when $(\hat{\mu}_n \ \hat{\sigma}_n)$ is a "reasonable" sequence of estimators of $(\mu \ \sigma)$. This provides us with quadratic nuisance parameter free goodness-of-fit tests for the location and scale parameters family of D.F.'s. These results are applied to the problems of testing for the exponential, the logistic, the Weibull and the extreme value distributions.

In chapter III, we obtain in-probability representation of the estimated quantile process. In addition, further goodness-of-fit considerations are also discussed.

Our chapter IV is concerned with exponential and uniform spacings. Instead of using the "quantile approach", we use the strong approximation results of Komlós, Major, Tusnády (1975, 1976). The results of Moran (1947), Kimball (1950), Koziol (1980), some results of Darling (1953) and some results of LeCam (1958) are special cases of our results. In addition, our general treatment enables us to introduce and discuss some new nuisance parameter free distribution free goodness-of-fit tests for the uniform and exponential distributions. These procedures can be applied to test for the random distribution of events in time. Moreover, some answers to the problem of rates of

convergence are given. The derived rates of convergence appear to be encouraging enough for statisticians to use our proposed test statistics and to start reusing some old tests, like for example that of Moran (1947).

In chapter V we consider the asymptotic distribution of a test for clustering in which the observations is partitioned into two groups so that the between groups sum of squares is maximized. The results of this chapter are generalization of the results of Brillinger, Knott, Scott (1979). In their paper they assumed that the observations are taken from a normal population, whereas in our treatment the normality assumption is not used. The results of this chapter can be looked at from the robustness point of view.

In this dissertation we wished to put emphasis not only on the particular results obtained, but also on the strong approximation methodology used throughout.

The results in chapters I, II, and III are joint with Professor M. Csörgő. Chapter IV is based on the author's paper ((1981), to appear). Chapter V is based on the author's paper ((1979), to appear). The results of all these chapters are believed to be new.

CHAPTER I

ON THE QUANTILE PROCESS

1. Introduction

Let $X_1, X_2, \ldots$ be a sequence of i.i.d.rv's with a continuous distribution function $F(.)$ and let $Y_1, Y_2, \ldots, Y_n$ denote the order statistics of a random sample $X_1, X_2, \ldots, X_n$. Define the empirical quantile function Q_n by.

$$Q_n(u) = Y_k \quad \text{if} \quad \frac{k-1}{n} < u \leq \frac{k}{n}, \quad k = 1,2,\ldots,n. \tag{1.1}$$

For $u = 0$, we define $Q_n(0) = Y_0 = Y_1$. It is piecewise constant function whose values are the order statistics $Y_1, Y_2, \ldots, Y_n$.

Parzen (1979 a,b) has suggested another definition of the empirical quantile function. In the sequel, we will call Parzen's definition the second defnition.

The second definition of $Q_n(u)$ is:

$$Q_n(u) = n(\frac{j}{n} - u)Y_{j-1} + n(u - \frac{j-1}{n})Y_j$$

$$\text{for} \quad \frac{j-1}{n} \leq u \leq \frac{j}{n} \quad \text{and} \quad j = 1,2,\ldots,n. \tag{1.2}$$

For $j = 1$, we define $Y_0 = Y_1$ and consequently the second definition of $Q_n(u)$ is

$$Q_n(u) = Y_1 \qquad \text{for } 0 \le u \le \frac{1}{n}$$

$$= n\left(\frac{j}{n} - u\right)Y_{j-1} + n\left(u - \frac{j-1}{n}\right)Y_j \qquad \text{for } \frac{j-1}{n} < u \le \frac{j}{n}$$

$$\text{and } j = 2,3,\ldots,n. \qquad (1.2)^*$$

It is a piecewise linear function. Then $\tilde{q}(u) = Q_n'(u)$ is given by;

$$\tilde{q}(u) = n(Y_j - Y_{j-1}) \qquad \text{for } \frac{j-1}{n} < u < \frac{j}{n} \text{ and } j = 1,2,\ldots,n. \qquad (1.3)$$

We begin with quoting results concerning strong approximation of the quantile process

$$q_n(u) = n^{1/2}(Q_n(u) - F^{-1}(u)) \quad , 0 < u < 1 \quad , \qquad (1.4)$$

with $Q_n(u)$ as in (1.1) and $F^{-1}(u) = \inf\{x : F(x) \ge u\}$. Define the standardized empirical quantile process ρ_n by

$$\rho_n(u) = f(F^{-1}(u))\, q_n(u) \quad , 0 < u < 1 \quad , \qquad (1.5)$$

with $q_n(u)$ as in (1.4).

<u>Theorem 1A</u> (Csörgö-Révész (1980) and (1978))

Let $X_1, X_2, \ldots$ be i.i.d.r.v with a continuous distribution function F which is also twice differentiable on (a,b), where $-\infty \le a = \sup\{x : F(x) = 0\}$, $+\infty > b = \inf\{x : F(x) = 1\}$ and $F' = f > 0$ on (a,b). Assume that for some $\gamma > 0$

$$\sup_{a<x<b} F(x)\,(1 - F(x))\,\frac{|f'(x)|}{f^2(x)} \le \gamma . \tag{1.6}$$

One can then define a Brownian bridge $\{B_n(u)\ ;\ 0 \le u \le 1\}$ for each n and a Kiefer process $\{K(y,t)\ ;\ 0 \le y \le 1\ ,\ 0 \le t\}$ such that

$$\sup_{\delta_n \le u \le 1-\delta_n} |\rho_n(u) - B_n(u)| \ \underline{a.s.}\ O(n^{-1/2}\log n) \tag{1.7}$$

and

$$\sup_{\delta_n \le u \le 1-\delta_n} |n^{1/2}\rho_n(u) - K(u,n)| \ \underline{a.s.}\ O((n\log\log n)^{1/4}(\log n)^{1/2}), \tag{1.8}$$

where $\delta_n = 25\, n^{-1}\log\log n$.

If, in addition to (1.6), we also assume that

$$A = \lim_{x\downarrow a} f(x) \quad , \quad B = \lim_{x\uparrow b} f(x) < \infty \tag{1.9}$$

with

(i) $\min(A,B) > 0$

or

(ii) if $A = 0$ (resp. $B = 0$), then f is nondecreasing (resp. nonincreasing) on an interval to the right of a (resp. to the left of b),

then (1.6) and (1.9) with (i) imply

$$\sup_{0 \le u < 1} |\rho_n(u) - B_n(u)| \ \underline{a.s.}\ O(n^{-1/2}\log n) \tag{1.10}$$

and (1.6) and (1.9) with (ii) imply

$$\sup_{0 \le u \le 1} |\rho_n(u) - B_n(u)|$$

$$\underline{a.s.}\ O(n^{-1/2}\log n) \qquad \text{if } \gamma < 2$$

$$\underline{a.s.}\ O(n^{-1/2}(\log\log n)^{\gamma}(\log n)^{(1+\varepsilon)(\gamma-1)}) \quad \text{if } \gamma \geq 2 \ , \tag{1.11}$$

where γ is as in (1.6) and $\varepsilon > 0$ is arbitrary; also (1.6) and (1.9) with (i) or (ii) imply

$$\sup_{0<u<1} |n^{1/2}\ \rho_n(u) - K(u,n)|\ \underline{a.s.}\ O((n\log\log n)^{1/4}(\log n)^{1/2}). \tag{1.12}$$

Remark 1.1

Throughout this thesis the notation $\underline{a.s.}\ O(.)$ as well as the notation $\overset{P}{=} O(.)$ is meant with an absolute constant.

As to the notion of a Brownian bridge we have

Definition 1.1

A Brownian bridge $\{B(y)\ ;\ 0 \leq y \leq 1\}$ is a separable Gaussian process with $EB(y) = 0$ and covariance function

$EB(y_1)\ B(y_2) = y_1 \wedge y_2 - y_1 y_2$.

We will also need

Definition 1.2

A Kiefer process $\{K(y,t)\ ;\ 0 \leq y \leq 1\ ,\ 0 \leq t\}$ is a separable Gaussian process with $EK(y,t) = 0$ and covariance function

$EK(y_1,t_1)\ K(y_2,t_2) = t_1 \wedge t_2 (y_1 \wedge y_2 - y_1 y_2)$.

The following is going to be useful in the sequel.

Theorem 1B (P. Lévy (1973), (1948)).

$$\lim_{h\to 0} (2h \log \frac{1}{h})^{-1/2} \sup_{0<s<1-h} |W(s+h) - W(s)| \text{ a.s. } 1 , \tag{1.13}$$

where $\{W(t); 0 \le t < \infty\}$ is a Wiener process, i.e., a continuous Gaussian process with $EW(t) = 0$ and covariance function $EW(t)\,W(s) = s \wedge t$.

2. Asymptotic Distribution Theorey for Parzen's Definition of Q_n

Parzen (1979 a,b) proposed a problem, which is to show that the above mentioned aymptotic distribution theory for the first definition of Q_n, i.e., as in (1.1) (cf. Theorem 1A), applies also to the second definition of Q_n. For the sake of tackling the latter we define

$$\lambda(n,u) = \{nu\} - nu , \tag{2.1}$$

where $\{x\}$ = the smallest integer greater than or equal to x, and

$$\begin{aligned} \rho_n^*(u) &= \sqrt{n}\, f(F^{-1}(u))(Y_1 - F^{-1}(u)) \qquad \text{for } 0 \le u \le \frac{1}{n} \\ &= \lambda(n,u)\sqrt{n}\, f(F^{-1}(u - \tfrac{1}{n}))\,[Y_{\{nu-1\}} - F^{-1}(u - \tfrac{1}{n})] \\ &\quad + (1 - \lambda(n,u))\sqrt{n}\, f(F^{-1}(u))[Y_{\{nu\}} - F^{-1}(u)] \\ &\qquad \text{for } \frac{1}{n} < u < 1 . \end{aligned} \tag{2.2}$$

Theorem 2.1

Let $X_1, X_2, \ldots$ be i.i.d.rv with a continuous distribution function F which is also twice differentiable on (a,b), where $-\infty \leq a = \sup\{x : F(x) = 0\}$, $+\infty \geq b = \inf\{x ; F(x) = 1\}$ and $F' = f \neq 0$ on (a,b). One can then define a Brownian bridge $\{B_n(u) ; 0 \leq u \leq 1\}$ for each n and a Kiefer process $\{K(y,t) ; 0 \leq y \leq 1 , t \geq 0\}$ such that if condition (1.6) and (1.9) with (i) are assumed, then

$$\sup_{0<u<1}|\rho_n^*(u) - B_n(u)| \overset{P}{=} O(n^{-1/2}\log n) \tag{2.3}$$

and if (1.6) and (1.9) with (ii) are assumed, then

$$\sup_{0<u<1}|\rho_n^*(u) - B_n(u)|$$

$$\overset{P}{=} O(n^{-1/2}\log n) \qquad \text{if } \gamma < 2$$

$$\overset{P}{=} O(n^{-1/2}(\log\log n)^{\gamma}(\log n)^{(1+\varepsilon)(\gamma-1)}) \text{ if } \gamma \geq 2 , \tag{2.4}$$

where γ as in (1.6) and $\varepsilon > 0$ is arbitrary. In addition if conditions (1.6) and (1.9) with (i) or (ii) are assumed, then

$$\sup_{0<u<1}|n^{-1/2}\rho_n^*(u) - K(u,n)| \underline{\text{a.s.}} O((n\log\log n)^{1/4}(\log n^{1/2})). \tag{2.5}$$

Proof.

We notice that the first definition of Q_n is

$$Q_n(u) = Y_{\{nu\}} .$$

For $\frac{1}{n} < u < 1$, we have

$$|\rho_n^*(u) - B_n(u)| = |\rho_n^*(u) - \lambda(n,u)\, B_n(u) - (1 - \lambda(n,u))\, B_n(u)|$$

$$= |\rho_n^*(u) - \lambda(n,u)\, B_n(u - \tfrac{1}{n}) - (1 - \lambda(n,u))\, B_n(u)$$

$$- \lambda(n,u)(B_n(u) - B_n(u - \tfrac{1}{n}))|$$

$$\leq \lambda(n,u)\, |\sqrt{n}\, f(F^{-1}(u - \tfrac{1}{n}))(Y_{\{nu-1\}} - F^{-1}(u - \tfrac{1}{n}))$$

$$- B_n(u - \tfrac{1}{n})|$$

$$+ (1 - \lambda(n,u))\, |\sqrt{n}\, f(F^{-1}(u))(Y_{\{nu\}} - F^{-1}(u)) - B_n(u)|$$

$$+ \lambda(n,u)\, |B_n(u) - B_n(u - \tfrac{1}{n})|$$

$$= \lambda(n,u)\, |\rho_n(u - \tfrac{1}{n}) - B_n(u - \tfrac{1}{n})|$$

$$+ (1 - \lambda(n,u))\, |\rho_n(u) - B_n(u)|$$

$$+ \lambda(n,u)\, |B_n(u) - B_n(u - \tfrac{1}{n})| \quad , \qquad (2.6)$$

with ρ_n is as in (1.5).

As to the third term, we apply Theorem 1B to $B(u) = W(u) - uW(1)$, $0 \leq u \leq 1$, and get

$$\sup_{\frac{1}{n}<u<1} |B_n(u) - B_n(\mu - \tfrac{1}{n})|$$

$$\overset{D}{=} \sup_{\frac{1}{n}<u<1} |B(u) - B(u - \tfrac{1}{n})|$$

$$\underline{\text{a.s.}} \quad O(n^{-1/2}(\log n)^{1/2}). \tag{2.7}$$

Consequently

$$\sup_{\frac{1}{n}<u<1} |B_n(u) - B_n(u - \tfrac{1}{n})| \overset{P}{=} O(n^{-1/2}(\log n)^{1/2}). \tag{2.8}$$

Thus by taking the supremum of both sides of (2.6) over $(\frac{1}{n}, 1)$ and taking into account (1.11), (1.12) and (2.8), we get

$$\sup_{\frac{1}{n}<u<1} |\rho_n^*(u) - B_n(u)| \overset{P}{=} O(n^{-1/2}\log n) \quad , \tag{2.9}$$

if (1.6) and (1.9) with (i) are assumed, and if (1.6) and (1.9) with (ii) are assumed, we get

$$\sup_{\frac{1}{n}<u<1} |\rho_n^*(u) - B_n(u)|$$

$$\overset{P}{=} O(n^{-1/2}\log n) \qquad \text{if } \gamma < 2$$

$$\overset{P}{=} O(n^{-1/2}(\log\log n)^{\gamma}(\log n)^{(1+\varepsilon)(\gamma-1)}) \text{ if } \gamma \geq 2 \ , \tag{2.10}$$

where γ is as in (1.6) and $\varepsilon > 0$ is arbitrary. Also,

$$\rho_n^*(u) = f(F^{-1}(u))\, q_n(u) = \rho_n(u) \quad \text{for} \quad 0 < u \le \frac{1}{n} , \tag{2.11}$$

with ρ_n is as in (1.5), and

$$\sup_{0<u<1} |\rho_n^*(u) - B_n(u)| \le \sup_{0<u\le \frac{1}{n}} |\rho_n^*(u) - B_n(u)|$$

$$+ \sup_{\frac{1}{n}<u<1} |\rho_n^*(u) - B_n(u)| \tag{2.12}$$

Now, if (1.6) and (1.9) with (i) are assumed, then (2.3) follows by (2.9) and by applying (1.10) to the first term of the right hand side of (2.12). On the other hand, if (1.6) and (1.9) with (ii) are assumed, then (2.4) follows by (2.10) and by applying (1.12) to the first term of the right hand of (2.12).

The statement of (2.5) can be proved similarly, using the inequality of (2.6), respectively that of (2.12), with $K(u,n)/n^{1/2}$ replacing $B_n(u)$ in it, and the fact that

$$\sup_{\frac{1}{n}<u<1} |K(u,n) - K(u - \tfrac{1}{n},n)| \ \underline{\text{a.s.}}\ O((\log n)^{1/2}) . \tag{2.13}$$

The latter in turn follows from

<u>Theorem 1C</u> (cf. Chan (1977) or Theo[illegible].15.2 in Csörgő and Révész (1981)).

Let $\{h_n\}$ be a sequence of positive numbers for which

$$\lim_{n\to\infty} \frac{\log h_n^{-1}}{\log \log n} = \infty$$

Then

$$\lim_{n\to\infty} \sup_{0\le u\le 1-h_n} \beta_n |K(u + h_n, n) - K(u,n)| \ \underline{\underline{a.s.}}\ 1$$

where $\beta_n = (2nh_n \log h_n^{-1})^{-1/2}$.

CHAPTER II

QUADRATIC NUISANCE PARAMETER FREE GOODNESS-OF-FIT TESTS IN THE PRESENCE OF LOCATION AND SCALE PARAMETERS

1. Introduction, Legend, Preliminaries and Definitions

Let $X_1, X_2, \ldots$ be a sequence of i.i.d.rv's with distribution function F with unknown location and scale parameters $-\infty < \mu < \infty$ and $\sigma > 0$ respectively, and assume that F is of the form $F(x;\mu,\sigma) = F_0((x-\mu)/\sigma)$, $x \in R$, where $F_0(z) = F(\sigma z + \mu;\mu,\sigma)$. Let $\mathcal{F}$ be the class of all continuous distribution functions of this latter form, i.e.,

$$\mathcal{F} = \{F(x;\mu,\sigma) : F(x;\mu,\sigma) = F_0(\frac{x-\mu}{\sigma}),\ -\infty < \mu < \infty,\ \sigma > 0\}. \qquad (1.1)$$

Further, let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution F. We are interested in testing the composite null hypothesis

$$H_0 : F \in \mathcal{F}\ ,\ \text{with}\ F_0\ \text{specified}, \qquad (1.2)$$

and $\mathcal{F}$ is as in (1.1).

Let F_n be the empirical distribution function of the random sample $X_1, X_2, \ldots, X_n$. The empirical process based on $X_1, X_2, \ldots, X_n$ is defined by

$$\beta_n(x) = n^{1/2}(F_n(x) - F(x;\mu,\sigma)) \quad , \quad -\infty < x < +\infty \quad , \tag{1.3}$$

and, as it is well known, its Kolmogorov-Smirnov functional $\sup_{-\infty<x<\infty} |\beta_n(x)|$ and its Cramér-von Mises functional $\int_{-\infty}^{\infty} \beta_n^2(x)\, dF(x)$ can be used to test the simple goodness-of-fit hypothesis that our random sample $X_1, X_2, \ldots, X_n$ has a completely specified distribution function, e.g.,

$$H_0' : F = F_0\left(\frac{x-\mu}{\sigma}\right) \quad , \quad \text{with } F_0 \text{ specified and } \mu, \sigma \text{ known.} \tag{1.4}$$

However, when testing for composite null hypothesis, like for example that of (1.2), we have the problem of not knowing the values of μ and σ. There are many ways of "getting rid of $\underset{\sim}{\theta} = (\mu\ \sigma)$" so as to reduce composite goodness-of-fit hypotheses to simple ones. As far as the empirical process is concerned, one natural way of doing this is to estimate θ by using some kind of "a good estimator" sequence $\{\hat{\theta}_n\}_{n=1}^{\infty}$, based on the sample $X_1, X_2, \ldots, X_n$ from $F(x;\theta)$, $\theta = (\mu\ \sigma)$.

Concerning the classical Cramér-von Mises and Kolmogorov-Smirnov statistics, Darling (1955) and Kac, Kiefer, Wolfowitz (1955) studied their asymptotic distributions, by first estimating the unknown parameters of some special continuous distribution functions. Durbin (1973a) studied the weak convergence of the empirical process under a given sequence of alternative hypotheses when parameters of any continuous distribution function $F(x;\theta)$ are estimated from the data via maximum likelihood or maximum likelihood-like estimators. Neuhaus (1974, 1976a,b) also considered the asymptotic properties of the Cramér-von Mises statistic when parameters are estimated. Burke

et al (1979) obtained asymptotic in-probability and almost sure representations in terms of Gaussian processes in both x and n for the empirical process when parameters are estimated by ML or ML-like estimators via the strong approximation methodology of Kiefer (1972), Csörgő, Révész (1975) and Komlós, Major, Tusnády (1975) for the empirical process β_n of (1.3).

A common drawback of the general results of the weak and strong convergence papers quoted in the above paragraph is that the resulting Gaussian processes which approximate the estimated empirical process depend, in general, on both the form of the hypothetical distribution function and the true values of the unknown parameters of the latter, resulting in a situation where the statistics under consideration are computable, given the form of F, but the distributions of the corresponding functionals of the approximating Gaussian process are not.

One way to overcome the above mentioned drawback is the so-called random substitution method suggested by Durbin (1961, 1976). In this method, the estimator $\hat{\theta}_n$ of the unknown θ calculated from the sample is to be replaced by a corresponding estimator $\hat{\theta}_n^*$, external to the sample, of a known value of θ. It is shown that for the family F of (1.1), under mild conditions, the limiting distribution of the thus estimated empirical process on the null hypothesis is the same as if the values of the nuisance parameters were known to begin with. Regretably, the random substitution method has not found favour in practice. The reasons seem to be, as suggested by Durbin (1976, p. 38),

(i) Practical workers do not like introducing into the analysis of real data the element of artificial randomization involved in selecting the random vector $\hat{\theta}_n^*$, external to the data.

(ii) It is suspected that a loss of power may result via the use of randomization.

(iii) The computational labor required to obtain $\hat{\theta}_n^*$ and effect the required transformation is burdensome.

A much simpler randomization way of constructing a form of the estimated empirical process which has the same limiting distribution as β_n under both null and alternative hypotheses is the so-called half-sample device. In this method the estimate of θ is calculated from a randomly chosen half of the given sample instead of from the whole sample. This device was suggested by Durbin (1973b) following up an earlier related proposal by Rao (1972). It was studied further by Durbin (1976), and Burke et al (1979). The main feature of this method is that on the null hypothesis the estimated empirical process has the same limiting distribution as if the values of the nuisance parameters were known. This suggests that the random substitution and random half-sample methods are asymptotically equivalent, i.e., the effect of the randomization element required to select the half-sample is asymptotically equivalent to that involved in the use of the random substitution method. However, the half-sample device is intuitively more appealing than the random substitution method, since the former does not require going outside the set of

observed values. In addition, it is computationally much easier to use in practice. Durbin (1976, p. 40) suggested that the half-sample device might provide a useful interim procedure in the absence or inconvenience of the full sample procedures.

Csörgő et al (1974) suggested another approximate solution, namely to put the estimator $\hat{\theta}_n$ in the limiting Gaussian process of the estimated empirical process, and they have proved, at least in principle, the applicability of this method.

As an alternative route to test for H_0 of (1.2), Csörgő, Révész (1980, I-II), following up a conjecture by Parzen (1979a,b), suggested tests based on the stochastic process

$$\{p_n(u) \; ; \; 0 \le u \le 1 \; , \; n = 1,2,\ldots\}$$

defined on the bases of a random sample $X_1, X_2, \ldots, X_n$ as

$$p_n(u) = \begin{cases} 0 & \text{if } 0 \le u < \frac{2}{n} \\ n^{1/2} \left\{ \dfrac{\sum_{1}^{[nu]-1} f_0(F_0^{-1}(\frac{j}{n+1}))(X_{j+1:n} - X_{j:n})}{\sum_{1}^{n-1} f_0(F_0^{-1}(\frac{j}{n+1}))(X_{j+1:n} - X_{j:n})} - u \right\} & \text{if } \frac{2}{n} \le u \le 1, \end{cases} \tag{1.5}$$

where $[x]$ stands for the integer part of the real number x and $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ are the order statistics of the sample $X_1, X_2, \ldots, X_n$. It is clear that, under H_0 the distribution of $p_n(u)$ is nuisance parameter free for each n and consequently its asymptotic

process is also nuisance parameter free. (Csörgő, Révész (1980I,II) derived the limiting Gaussian process which turned out to be a Brownian bridge in the uniform and exponential cases but otherwise it depends on the form of F_0.

Another alternative route to test for H_0 of (1.2), is to consider tests based on the stochastic process

$$\{g_n(u;\alpha)\ ;\ 0 \le u \le 1\ ,\ n = 1,2,\ldots,\ \alpha > 0\}$$

defined on the bases of a random sample $X_1, X_2, \ldots, X_n$ as follows;

$$g_n(u;\alpha) = \begin{cases} 0 & \text{if } 0 \le u < \frac{2}{n} \\ n^{1/2}\left[\dfrac{n^{-1}\sum_{1}^{[nu]-1} f_0^{\alpha}(F_0^{-1}(\frac{j}{n+1}))(X_{j+1:n} - X_{j:n})^{\alpha}}{(\sum_{1}^{n-1} f_0(F_0^{-1}(\frac{j}{n+1}))(X_{j+1:n} - X_{j:n}))^{\alpha}} - u\Gamma(1+\alpha)\right] & \text{if } \frac{2}{n} \le u \le 1 . \end{cases} \tag{1.6}$$

Again, given H_0 of (1.2) the distribution of $g_n(u;\alpha)$ is nuisance parameter free for each n and consequently its asymptotic process is also nuisance parameter free and it is excepted that its asymptotic process will depend on the form of F_0. However, it appears that the asymptotic theory of $g_n(u;\alpha)$ is extremely difficult in general even for the important case of $\alpha = 2$, i.e., for the L^2-version of $p_n(u)$ of (1.5) of Csörgő, Révész.

Concerning the uniform and exponential cases, we studied the asymptotic theory of $g_n(u;\alpha)$ of (1.6) in chapter IV. We proved that, just like that of $p_n(u)$ of (1.5), the limiting Gaussian process is nuisance parameter free and distribution free in the latter two cases.

For the sake of further discussion and also for later use, we define now the empirical quantile function Q_n by

$$Q_n(y) = \begin{cases} X_{k:n} & \text{if } \frac{k-1}{n+1} < y \leq \frac{k}{n+1}, \quad k = 1,2,\ldots,n \\ X_{n:n} & \text{if } \frac{n}{n+1} < y < 1 \end{cases} , \tag{1.7}$$

where $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ are the order statistics of the sample $X_1, X_2, \ldots, X_n$, and if the continuous distribution function F of the sample has a density function f, the standardized empirical quantile process ρ_n by

$$\rho_n(u) = n^{1/2} f(F^{-1}(u))(Q_n(u) - F^{-1}(u)) \quad , 0 < u < 1 , \tag{1.8}$$

where $F^{-1}(u) = \inf\{x : F(x) \geq u\}$, the inverse of F.

From now on, unless explicitly indicated otherwise, let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution function F and let $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ be the corresponding order statistics.

Our task is to test the composite goodness-of-fit hypothesis of (1.2), namely, $H_0 : F \in \mathcal{F}$ of (1.1) with F_0 specified.

Let us assume that F_0 of (1.1) has a density function f_0, and define the standardized empirical quantile process ρ_n^0 of the family $\mathcal{F}$ of (1.1) by

$$\rho_n^0(u) = \rho(u;\mu,\sigma) = n^{1/2} f_0(F_0^{-1}(u))((Q_n(u) - \mu)/\sigma - F_0^{-1}(u))$$

$$= n^{1/2} f_0(F_0^{-1}(u))(Q_n^0(u) - F_0^{-1}(u)) , \quad 0 < u < 1 , \quad (1.9)$$

where $Q_n^0(u) = (Q_n(u) - \mu)/\sigma$ and $Q_n(u)$ is as in (1.7).

Let $\hat{\theta}_n = (\hat{\mu}_n \ \hat{\sigma}_n)$ be a sequence of estimators for $\theta = (\mu \ \sigma)$ and define the estimated standardized empirical quantile process $\hat{\rho}_n^0(u)$ as

$$\hat{\rho}_n^0(u) = \rho_n(u; \hat{\mu}_n, \hat{\sigma}_n)$$

$$= n^{1/2} f_0(F_0^{-1}(u)) \{(Q_n(u) - \hat{\mu}_n)/\hat{\sigma}_n - F_0^{-1}(u)\}, \quad 0 < u < 1 . \quad (1.10)$$

We return now to our preliminary discussion. Shapiro and Wilk (1965), and Shapiro and Francia (1972) have proposed a test statistic for normality (parameters unknown) which is perhaps the most accepted and successful one so far. De Wet and Venter (1972) have shown that the Shapiro and Francia test for normality, which itself is a large sample version of the original Shapiro-Wilk procedure, is asymptotically equivalent to rejecting the null hypothesis that the data have come from a normal distribution for large values of the statistic L_n' ,

which in our notation is defined as

$$L_n' = \sum_{k=1}^{n} \frac{(\hat{\rho}_n^0 (\frac{k}{n+1}))^2}{n\,\phi^2(\Phi^{-1}(\frac{k}{n+1}))} , \tag{1.11}$$

where, for the normal family, $\hat{\rho}_n^0$ is defined as.

$$\hat{\rho}_n^0(u) = n^{1/2}\,\phi(\Phi^{-1}(u))\left(\frac{Q_n(u) - \hat{\mu}_n}{\hat{\sigma}_n} - \Phi^{-1}(u)\right) , \quad 0 < u < 1 , \tag{1.12}$$

and in (1.11), $\hat{\theta}_n = (\hat{\mu}_n \quad \hat{\sigma}_n) = (\bar{X}_n \quad S_n)$, $\bar{X}_n$ is the sample mean, S_n^2 is the sample variance, ϕ and Φ , respectively are the density function and distribution function of a standard normal rv. Assuming normality, De Wet and Venter (1972) showed that

$$L_n = (L_n' - a_n) \xrightarrow{D} \sum_{k=3}^{\infty} (Z_k^2 - 1)/k ; \tag{1.13}$$

where $Z_1, Z_2, \ldots$ are i.i.d. standard normal rv, $a_n \simeq EL_n'$, whose approximate values they also tabulated, and they provided tables for the asymptotic distribution of L_n .

De Wet and Venter (1973) also treated H_0 of (1.2) with $\mu = 0$ and $\sigma > 0$. Along the lines of the original Shapiro-Wilk (1965) procedure, they proposed the test statistic E_n , defined by

$$E_n = \sum_{k=1}^{n} \frac{(\hat{\rho}_n^0(\frac{k}{n+1}))^2}{n\,f_0^2(F_0^{-1}(\frac{k}{n+1}))} \cdot \frac{J(\frac{k}{n+1})}{F_0^{-1}(\frac{k}{n+1})} , \tag{1.14}$$

where, $J(u) = \frac{1}{I}\, L'(F_0^{-1}(u))$, with $L(y) = -1 - y\,\frac{f_0'(y)}{f_0(y)}$, and

$I = \int_0^1 L'(F_0^{-1}(u))\, F_0^{-1}(u)\, du.$, and

$$\hat{\rho}_n^0\left(\frac{k}{n+1}\right) = \sqrt{n}\, f_0\left(F_0^{-1}\left(\frac{k}{n+1}\right)\right)\left(Q_n\left(\frac{k}{n+1}\right)/\hat{\sigma}_n - F_0^{-1}\left(\frac{k}{n+1}\right)\right)$$

$$\hat{\sigma}_n = \frac{1}{n}\sum_{k=1}^{n} J\left(\frac{k}{n+1}\right) X_{k:n} .$$

Letting $e_n \simeq E\,E_n$, they have found the limiting distribution of

$$E_n - e_n \quad . \tag{1.15}$$

Csörgő and Révész (1979), considering the same composite goodness-of-fit hypothesis of (1.2), aiming at such statistics which would not require norming factors, like for example a_n , e_n of L_n and E_n respectively, for their limits to exist, and which would still have a tail-sensitivity, though less pronounced than that of L_n and E_n , have proposed the following test statistic

$$M_n(\lambda) = \sum_{k=1}^{n} \frac{\{\hat{\rho}_n^0(\frac{k}{n+1})\}^2}{n\, f_0(F_0^{-1}(\frac{k}{n+1}))} \left(F_0^{-1}\left(\frac{k}{n+1}\right)\right)^{\lambda-1} , \tag{1.16}$$

where $\lambda \geq 1$ is a fixed integer, $\hat{\rho}_n^0$ is as in (1.10). Considering those f_0 for which $\int_{-\infty}^{+\infty} x\, f_0(x)dx = 0$ and $\int_{-\infty}^{+\infty} x^2\, f_0(x)dx = 1$ and taking $\hat{\theta}_n = (\hat{\mu}_n \;\; \hat{\sigma}_n) = (\bar{X}_n \;\; S_n)$ they derived the asymptotic distribution of $M_n(\lambda)$, which turned out to be nuisance parameter free, via that of

$$M_n^0(\lambda) = \sum_{k=1}^{n} \frac{\{\rho_n^0(\frac{k}{n+1})\}^2}{n\, f_0(F_0^{-1}(\frac{k}{n+1}))} (F_0^{-1}(\frac{k}{n+1}))^{\lambda-1} , \tag{1.17}$$

where ρ_n^0 is as in (1.9). They have proved the following Theorem.

<u>Theorem 2A</u> (Csörgő, Révész (1979), (1981)).

Let $X_1, X_2, \ldots, X_n$ be a random sample with an absolutely continuous distribution F_0, which satisfies all the conditions of Theorem 1A. In addition assume also that

$$\lim_{y\to 0} y^{\frac{1}{r}} |F_0^{-1}(y)| = \lim_{y\to 1} (1-y)^{\frac{1}{r}} |F_0^{-1}(y)| = 0 , \tag{1.18}$$

for some $r > 2\lambda$. Then via the sequence of Brownian bridges of Theorem 1A, we have

$$|M_n^0(\lambda) - n^{-1} \sum_{k=1}^{n} \{B_n^2(\frac{k}{n+1})/f_0(F_0^{-1}(\frac{k}{n+1}))\} (F_0^{-1}(\frac{k}{n+1}))^{\lambda-1}|$$

$$\underline{\underline{a.s.}}\ o(1) , \tag{1.19}$$

and

$$\left| M_n^0(\lambda) - \int_0^1 B_n^2(y)\, \lambda^{-1}\, d(F_0^{-1}(y))^{\lambda} \right| \overset{P}{=} o(1) , \lambda = 1,2,\ldots . \tag{1.20}$$

Remark 1.1

In the same quoted paper of Csörgő, Révész, they proved that

$$M^0(\lambda) = \int_0^1 B^2(y)\ \lambda^{-1}\ d(F_0^{-1}(y))^{\lambda} \tag{1.21}$$

exists a.s.

Csörgő, Révész((1979), (1981)) pointed out that it was desirable to work out an analogue of their results about $M_n(\lambda)$, taking the route of having more general estimators $(\hat{\mu}_n \quad \hat{\sigma}_n)$, other than $(\bar{X}_n \quad S_n)$, which they used.

One of the aims of this chapter is to address to this general route. Indeed, we show that a similar asymptotic theory holds true for any other "reasonable" method of estimation, like for example maximum likelihood and maximum likelihood like estimators, L-estimators, or for the Weiss' estimate for σ when μ is known.

We mention here that the results of section 3 are, in addition to being essential for those of section 4, of special interest in themselves. More precisely, we show that under suitable restrictions we have uniformly in $0 \le u < v \le 1$,

$$n^{1/2}\left(\frac{1}{n}\sum_{[nu]+1}^{[nv]} J\left(\frac{k}{n+1}\right) X_{k:n} - \int_u^v J(x)\ G^{-1}(x)dx\right)$$

$$\xrightarrow{D} \int_u^v J(x)\ B(x)\ dG^{-1}(x) \tag{1.22}$$

where $J(.)$ is a well behaved function, $B(.)$ is a Brownian bridge and

$X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ are the order statistics of a random sample $X_1, X_2, \ldots, X_n$ with a distribution function $G(.)$.

Continuing with the preliminaries, for the sample $X_1, X_2, \ldots, X_n$ from $F \in \mathcal{F}$ of (1.1) and under H_0 of (1.2), define $U_i = F_0((X_i - \mu)/\sigma)$, $i = 1, 2, \ldots, n$. Then $U_1, U_2, \ldots, U_n$ are i.i.d.rv's with uniform distribution on $[0,1]$. Construct the uniform empirical process $\alpha_n(.)$ and the uniform empirical quantile process $u_n(.)$, of the sample $U_1, U_2, \ldots, U_n$, the following way:

$$\alpha_n(y) = n^{1/2}(F_n(y) - y) \quad , 0 < y < 1 \quad , \tag{1.23}$$

and

$$u_n(y) = n^{1/2}(Q_n(y) - y) \quad , 0 < y < 1 \quad , \tag{1.24}$$

where $F_n(.)$ and $Q_n(.)$ are the empirical distribution function and the empirical quantile function, based on $U_1, U_2, \ldots, U_n$, respectively.

Komlós, Major, Tusnády (1975) proved that there exists a sequence of Brownian bridges $\{B_n(y), 0 \leq y \leq 1\}$ for each n such that

$$\sup_{0 \leq y \leq 1} |\alpha_n(y) - B_n(y)| \underline{\underline{a.s.}} O(n^{-1/2} \log n). \tag{1.25}$$

Kiefer (1970) proved that

$$\sup_{0 \leq y \leq 1} |\alpha_n(y) - (-u_n(y))| \underline{\underline{a.s.}} O((n^{-1} \log \log n)^{1/4} (\log n)^{1/2}). \tag{1.26}$$

Combining (1.25) and (1.26), we get

$$\sup_{0 \leq y \leq 1} |u_n(y) - B_n(y)| \underline{\underline{a.s.}} O((n^{-1} \log \log n)^{1/4} (\log n)^{1/2}) . \tag{1.27}$$

Csörgő, Révész (1978) proved that, if F_0 satisfies all the conditions of Theorem 1A, then

$$\sup_{0<y<1}|\rho_n^0(y) - u_n(y)|$$

$$\underline{\underline{a.s.}} \begin{cases} O(n^{-1/2}\log\log n) & \gamma < 1 \\ O(n^{-1/2}(\log\log n)^2) & \gamma = 1 \\ O(n^{-1/2}(\log\log n)^{\gamma}(\log n)^{(1+\varepsilon)(\gamma-1)}) \text{ if } & \gamma > 1 , \end{cases} \tag{1.28}$$

where γ is as in (1.6) of chapter I and $\varepsilon > 0$ is arbitrary.

Combining now (1.27) and (1.28), we get that under all the conditions of Theorem 1A

$$\sup_{0<y<1} |\rho_n^0(y) - B_n(y)| \ \underline{\underline{a.s.}}\ O((n^{-1}\log\log n)^{1/4}(\log n)^{1/2}). \tag{1.29}$$

In summarizing the above few lines, we have

<u>Corollary 1.1</u>

Given all the conditions of Theorem 1A concerning F_0 of $\rho_n^0(.)$ of (1.9) and letting $\alpha_n(.)$ be as in (1.23), with the Brownian bridges $\{B_n\}$ of (1.25) we have

$$\sup_{0\le y\le 1} |\sqrt{n}(F_n(F^{-1}(y)) - y) - B_n(y)|$$

$$= \sup_{0\le y\le 1} |\alpha_n(y) - B_n(y)| \ \underline{\underline{a.s.}}\ O(n^{-1/2}\log n) \tag{1.30}$$

and

$$\sup_{0<y<1} |\rho_n^0(y) - B_n(y)| \ \underline{a.s.}\ O((n^{-1}\log\log n)^{1/4}(\log n)^{1/2}). \qquad (1.31)$$

Remark 1.1

A clear disadvantage of corollary (1.1) is that the rate of (1.31) is far from the best possible one, which is given by Csörgő, Révész (1978). However, with respect to our present needs (i.e., to prove Theorem 2.1) corollary (1.1) has the advantage of giving a representation for the processes $\alpha_n(y)$ and $\rho_n^0(y)$ in terms of the same sequence of Brownian bridges, and the shown rate of convergence in (1.31) is sufficient for our purpose here.

2. The Asymptotic Distribution of $M_n(\lambda)$ when Using M.L. and M.L-Like Estimators

For an i.i.d. sequence $X_1, X_2, \ldots$ from the family of distribution functions F of (1.1), let $\hat{\theta}_n = (\hat{\mu}_n \ \hat{\sigma}_n)$ be a sequence of estimators of $\theta = (\mu\ \sigma)$ based on $X_1, X_2, \ldots, X_n$. Adapting the maximum likelihood like conditions of Burke et al (1979) for the special parameter structure of the family F, we assume:

(i) $n^{1/2}(\theta_n - \theta_0) = n^{-1/2} \sum_1^n \ell(X_j,\theta_0) + \varepsilon_n$, where $\ell(.,\theta_0)$ is a measurable 2-dimensional vector valued function, $\theta_0 = (\mu_0 \ \sigma_0)$ is the theoretical value of $\theta = (\mu\ \sigma)$, and $\varepsilon_n \xrightarrow{P} 0$.

(ii) $E\,\ell(X_j, \theta_0) = 0$, $j = 1, 2, \ldots, n$. (2.1)

(iii) $E\{\ell(X_j,\theta_0)^t\,\ell(X_j,\theta_0)\}$ is a finite non-negative definite matrix, where A^t is the transpose of the matrix A.

(iv) Each component of the vector function $\ell(x,\theta_0)$ is of bounded variation on each finite interval.

(v) $\lim_{h\downarrow 0}\left(h \log\log\frac{1}{h}\right)^{1/2} \| \ell(F^{-1}(h,\theta_0),\theta_0)\| = 0$

and

$$\lim_{h\uparrow 1}\left((1-h)\log\log\frac{1}{1-h}\right)^{1/2}\|\ell(F^{-1}(h,\theta_0),\theta_0)\| = 0$$

where $\|\ \|$ on R^2 is defined by $\|(y_1,y_2)\| = \max(|y_1|,|y_2|)$.

Further assume that

$$\begin{aligned}\ell(x,\theta_0) &= (\ell_1(x,\theta_0) \quad \ell_2(x,\theta_0))^t \\ &= \left(\sigma_0\ell_1\left(\frac{x-\mu_0}{\sigma_0}\right) \quad \sigma_0\ell_2\left(\frac{x-\mu_0}{\sigma_0}\right)\right)^t \qquad (2.2)\end{aligned}$$

<u>Remark 2.1</u>

We note that condition (2.2) is natural enough in the sense that when both μ and σ are unknown (cf., e.g., Pitman (1939)) $\hat{\mu}_n$ and $\hat{\sigma}_n$ should satisfy

$$\left.\begin{aligned}\hat{\mu}_n(\gamma^{-1}(\underset{\sim}{x}+\underset{\sim}{\lambda})) &= \gamma^{-1}(\hat{\mu}_n(\underset{\sim}{x})+\lambda)\\ \hat{\sigma}_n(\gamma^{-1}(\underset{\sim}{x}+\underset{\sim}{\lambda})) &= \gamma^{-1}\hat{\sigma}_n(\underset{\sim}{x})\end{aligned}\right\} \qquad (2.3)$$

where $\underset{\sim}{x} = (x_1 \quad x_2 \quad \ldots \quad x_n)$, $\underset{\sim}{\lambda} = (\lambda \quad \lambda \quad \ldots \quad \lambda)$ and $\lambda, \gamma \in R$.

Now, we state and prove the following

Lemma 2.1

Assuming that conditions (2.1) (i)-(v) and (2.2) are satisfied, we have

$$n^{1/2}(\hat{\mu}_n - \mu_0) = \sigma_0 \int \ell_1(x)\, d_x\, B_n(F_0(x)) + \varepsilon^*_{1n} \tag{2.4}$$

and

$$n^{1/2}(\hat{\sigma}_n - \sigma_0) = \sigma_0 \int \ell_2(x)\, d_x\, B_n(F_0(x)) + \varepsilon^*_{2n} \quad , \tag{2.5}$$

where $\varepsilon^*_{1n}, \varepsilon^*_{2n} \xrightarrow{P} 0$ and the sequence $\{B_n(x) \; ; \; 0 \leq x \leq 1\}$ of Brownian bridges is the same sequence that has been used to approximate the empirical process of (1.23).

Proof

Consider

$$n^{1/2}(\hat{\mu}_n - \mu_0) = n^{-1/2} \sum_{j=1}^{n} \ell_1(X_j, \theta_0) + \varepsilon_{1n}$$

$$= \int \ell_1(x, \theta_0)\, d_x\, n^{1/2}\, F^*_n(x) + \varepsilon_{1n} \quad , \tag{2.6}$$

where $F^*_n(.)$ is the empirical distribution function based on $X_1, X_2, \ldots, X_n$.

Now, by (2.1) (ii), we have

$$n^{1/2}(\hat{\mu}_n - \mu_0) = \int \ell_1(x,\theta_0)\, d_x\, n^{1/2}(F_n^*(x) - F(x))$$

$$+ n^{1/2} \int \ell_1(x,\theta_0)\, d_x\, F(x) + \varepsilon_{1n}$$

$$= \int \ell_1(x,\theta_0)\, d_x\, \alpha_n(F(x)) + \varepsilon_{1n}$$

$$= \int \ell_1(x,\theta_0)\, d_x(\alpha_n(F(x)) - B_n(F(x)))$$

$$+ \int \ell_1(x,\theta_0)\, d_x\, B_n(F(x)) + \varepsilon_{1n} \quad . \tag{2.7}$$

Burke et al (1979) proved that if conditions (2.1) (iii) and (iv) are satisfied, we have

$$\int \ell_1(x,\theta_0)\, d_x(\alpha_n(x) - B_n(F(x)) \xrightarrow{P} 0 \quad .$$

Consequently, we have

$$n^{1/2}(\hat{\mu}_n - \mu_0) = \int \ell_1(x,\theta_0)\, d_x\, B_n(F(x)) + \varepsilon_{1n}^*$$

$$= \int \ell_1(F^{-1}(x;\mu_0,\sigma_0)\ ;\ \theta_0)\, d_x\, B_n(x) + \varepsilon_{1n}^*$$

$$= \int \ell_1(\mu_0 + \sigma_0 F_0^{-1}(x) \; ; \; \theta_0) \, d_x \, B_n(x) + \varepsilon_{1n}^*$$

$$= \sigma_0 \int \ell_1(F_0^{-1}(x)) \, d_x \, B_n(x) + \varepsilon_{1n}^* \quad , \qquad (2.8)$$

where $\varepsilon_{1n}^* \xrightarrow{P} 0$ and the last two lines of (2.9) are true since $F^{-1}(x) = \mu_0 + \sigma_0 F_0^{-1}(x)$ and $\ell_1(x;\theta_0) = \sigma_0 \ell_1(\frac{x - \mu_0}{\sigma_0})$ by (2.2). This proves (2.4). The proof of (2.5) is similar.

Corollary 2.1

Under the conditions of Theorem 1A and Lemma 2.1, the estimated standardized empirical quantile process $\hat{\rho}_n^0(y)$ (cf. (1.10)) can be written in terms of the standardized empirical quantile process $\rho_n^0(y)$ as follows.

$$\hat{\rho}_n^0(y) = \frac{1}{1 + o_p(1)} \{\rho_n^0(y) - T_n^{(1)} \, f_0(F_0^{-1}(y))$$

$$- T_n^{(2)} \, f_0(F_0^{-1}(y)) . F_0^{-1}(y) + o_p(1)(f_0(F_0^{-1}(y))$$

$$+ f_0(F_0^{-1}(y)) F_0^{-1}(y))\} . \qquad (2.9)$$

Proof

By (2.4) and (2.5), we have

$$\hat{\rho}_n^0(y) = n^{1/2} \, f_0(F_0^{-1}(y)) \left\{ \frac{Q_n(y) - \hat{\mu}_n}{\hat{\sigma}_n} - F_0^{-1}(y) \right\}$$

$$= \frac{1}{1 + (\frac{\hat{\sigma}_n}{\sigma_0} - 1)} \left\{ \rho_n^0(y) - \frac{\hat{\mu}_n - \mu_0}{\sigma_0} n^{1/2} f_0(F_0^{-1}(y)) \right.$$

$$\left. - \frac{\hat{\sigma}_n - \sigma_0}{\sigma_0} n^{1/2} f_0(F_0^{-1}(y)) F_0^{-1}(y) \right\}$$

$$= \frac{1}{1 + n^{-1/2} T_n^{(1)} + o_p(n^{-1/2})} \left\{ \rho_n^0(y) - (T_n^{(1)} + o_p(1)) f_0(F_0^{-1}(y)) \right.$$

$$\left. - (T_n^{(2)} + o_p(1)) f_0(F_0^{-1}(y)) F_0^{-1}(y) \right\} , \qquad (2.9)$$

where

$$T_n^{(i)} = \int \ell_i(x) \, d_x B_n(F_0(x)) \qquad , i = 1,2. \qquad (2.10)$$

This proves (2.9).

At this stage we are able to state and prove our main result of this section, namely;

<u>Theorem 2.1</u>

Let $X_1, X_2, \ldots, X_n$ be a random sample with a distribution function $F \in \mathcal{F}$ of (1.1). Let $\hat{\theta}_n = (\hat{\mu}_n \quad \hat{\sigma}_n)$ be a sequence of estimators of $\theta = (\mu \quad \sigma)$ based on $X_1, X_2, \ldots, X_n$ such that

conditions (2.1) (i)-(v) and (2.2) are satisfied. Further assume also that F_0 satisfies all the conditions of Theorem 2A and that

$$\int_{-\infty}^{\infty} f_0^3(x)\, dx < +\infty \tag{2.11}$$

Then, there exists a sequence of Brownian bridges $\{B_n\}$ such that for $\lambda = 1, 2, \ldots$ we have

$$|M_n^{(1)}(\lambda) - G_n^{(1)}(\lambda)| \xrightarrow{P} 0 \; ; \tag{2.12}$$

where $M_n^{(1)}(\lambda)$ is the ML or ML-like version of $M_n(\lambda)$ of (1.16), and

$$\begin{aligned} G_n^{(1)}(\lambda) = & \int_0^1 B_n^2(y)\lambda^{-1}\, d(F_0^{-1}(y))^{\lambda} \\ & + (T_n^{(1)})^2\, J^{(\lambda-1)} + (T_n^{(2)})^2\, J^{(\lambda+1)} \\ & - 2(T_n^{(1)}R_n^{(\lambda-1)} + T_n^{(2)}R_n^{(\lambda)} - T_n^{(1)}T_n^{(2)}J^{(\lambda)}) \; , \end{aligned} \tag{2.13}$$

and where

$$\begin{aligned} J^{(\alpha)} &= \int_0^1 f_0(F_0^{-1}(y))\, (F_0^{-1}(y))^{\alpha}\, dy \\ &= \int_{-\infty}^{+\infty} x^{\alpha} f_0^{\,2}(x)\, dx \qquad , \; \alpha = \lambda-1,\ \lambda,\ \lambda+1 \; , \end{aligned} \tag{2.14}$$

$$R_n^{(\alpha)} = \int_0^1 B_n(y)\ (F_Q^{-1}(y))^{\alpha}\ dy \qquad \alpha = \lambda-1,\ \lambda, \tag{2.15}$$

and $T_n^{(1)}$ and $T_n^{(2)}$ are as defined in (2.10).

<u>Proof</u>

By corollary (2.1), we have

$$M_n^{(1)}(\lambda) = \sum_{k=1}^{n} \left\{ n^{-1}\ (\hat{\rho}_n^0(\tfrac{k}{n+1}))^2 / f_0(F_0^{-1}(\tfrac{k}{n+1})) \right\} (F_0^{-1}(\tfrac{k}{n+1}))^{\lambda-1}$$

$$= \frac{1}{1+o_p(1)} \left\{ M_n^0(\lambda) + (T_n^{(1)})^2\ J_n^{(\lambda-1)} + (T_n^{(2)})^2\ J_n^{(\lambda+1)} \right.$$

$$\left. - 2 \left[T_n^{(1)}\ R_{n,\lambda-1} + T_n^{(2)}\ R_{n,\lambda} - T_n^{(1)}\ T_n^{(2)}\ J_n^{(\lambda)} \right] \right\}$$

$$+ o_p(1) \left\{ J_n^{(\lambda-1)} + 2J_n^{(\lambda)} + J_n^{(\lambda+1)} \right\}$$

$$+ 2\ o_p(1) \left\{ R_{n,\lambda-1} - T_n^{(1)}\ J_n^{(\lambda-1)} - T_n^{(2)}\ J_n^{(\lambda)} + R_{n,\lambda} \right.$$

$$\left. - T_n^{(1)}\ J_n^{(\lambda)} - T_n^{(2)}\ J_n^{(\lambda+1)} \right\}, \tag{2.16}$$

where $M_n^0(\lambda)$ is as in (1.17), $T_n^{(i)}$, $i = 1,2$ are as in (2.10), and

$$J_n^{(\alpha)} = n^{-1} \sum_{k=1}^{n} f_0(F_0^{-1}(\frac{k}{n+1}))\ (F_0^{-1}(\frac{k}{n+1}))^{\alpha}$$

$$, \quad \alpha = \lambda-1,\ \lambda,\ \lambda+1 \tag{2.17}$$

$$R_{n,\alpha} = n^{-1} \sum_{k=1}^{n} \rho_n^0\ (\frac{k}{n+1})\ (F_0^{-1}(\frac{k}{n+1}))^{\alpha}$$

$$, \quad \alpha = \lambda-1,\ \lambda\ . \tag{2.18}$$

Now (cf. Csörgő, Révész (1979), (1981))

$$J_n^{(\alpha)} \to J^{(\alpha)} \quad \text{as } n \to \infty \quad , \alpha = \lambda-1,\ \lambda,\ \lambda+1 \tag{2.19}$$

and

$$|R_{n,\alpha} - R_n^{(\alpha)}| = 0_p(1) \qquad , \alpha = \lambda-1,\ \lambda\ , \tag{2.20}$$

where $J^{(\alpha)}$ and $R_n^{(\alpha)}$ are as in (2.14) and (2.15) respectively. This together with Lemma 2.1, implies that the terms on the last three lines of (2.16) converge to zero in probability as $n \to \infty$, for $\lambda = 1, 2, \ldots$. Therefore, by Theorem 2A the required result follows.

Remark 2.1

Csörgő, Révész (1979), (1981) proved that $R_n^{(\alpha)}$ exists with probability one provided that all the conditions of Theorem 2A are

satisfied. In addition conditions (2.1) (iv) and (v) ensures the almost sure existence of $T_n^{(1)}$ and $T_n^{(2)}$. By (2.2) with $r > 2(\lambda+1)$ and (2.11) combined with Schwarz's inequality, it follows that $J^{(\alpha)}$, $\alpha = \lambda-1, \lambda, \lambda+1$ exist.

Corollary 2.2

The conditions of Theorem 2.1 imply

$$M_n^{(1)}(\lambda) \xrightarrow{D} G^{(1)}(\lambda) \quad , \quad \lambda = 1, 2, \ldots \quad , \tag{2.21}$$

where

$$G^{(1)}(\lambda) = \int_0^1 B^2(y)\lambda^{-1}\, d(F_0^{-1}(y))^{\lambda} + (T^{(1)})^2\, J^{(\lambda-1)} + (T^{(2)})^2\, J^{(\lambda+1)}$$

$$- 2\left\{ T^{(1)}\, R^{(\lambda-1)} + T^{(2)}\, R^{(\lambda)} - T^{(1)}T^{(2)}J^{(\lambda)}\right\} , \tag{2.22}$$

$$T^{(i)} = \int_0^1 \ell_i\, (F_0^{-1}(x))\, dB(x) \quad , \quad i = 1,2 \quad , \tag{2.23}$$

$$R^{(\alpha)} = \int_0^1 B(y)\, (F_0^{-1}(y))^{\alpha}\, dy \quad , \quad \alpha = \lambda-1, \lambda \tag{2.24}$$

and $B(.)$ is a Brownian bridge.

Corollary 2.3

If μ is unknown and σ is known, we have

$$M_n^{(1)}(\lambda) \xrightarrow{D} \int_0^1 B^2(y)\lambda^{-1}\, d(F_0^{-1}(y))^{\lambda} + (T^{(1)})^2 J^{(\lambda-1)} - 2T^{(1)} R^{(\lambda-1)} , \tag{2.25}$$

while, if μ is known and σ is unknown, then we have

$$M_n^{(1)}(\lambda) \xrightarrow{D} \int_0^1 B^2(y)\lambda^{-1}\, d(F_0^{-1}(y))^{\lambda} + (T^{(2)})^2 J^{(\lambda+1)} - 2T^{(2)} R^{(\lambda)} . \tag{2.26}$$

3. Linear Combinations of Order Statistics

Let $X_1, X_2, \ldots, X_n$ be i.i.d.rv's with continuous distribution function G. Let $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ be the corresponding order statistics. Consider statistics of the form

$$T_n(u,v) = n^{-1} \sum_{[nu]+1}^{[nv]} J\left(\frac{j}{n+1}\right) X_{j:n} , \quad 0 \le u < v \le 1 , \tag{3.1}$$

where $J(.)$ is a well behaved function. The latter class of statistics has received considerable attention in recent years. It has been shown, under suitable restrictions, that

$$n^{1/2}\left[T_n - \int_0^1 J(x)\, G^{-1}(x)\, dx\right] \xrightarrow{D} N(0,\sigma^2) \quad , \tag{3.2}$$

where $T_n = T_n(0,1)$ and

$$\sigma^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J(G(x))\, J(G(y))\, (G(\min(x,y)) - G(x)\, G(y))\, dxdy \ . \tag{3.3}$$

(cf., for example, Chernoff, Gastwirth and Johns (1967), Moore (1968) and Stigler (1974), (1979)).

The aim of the present section is two-fold. First, we wish to construct a "quantile representation" of the statistics $T_n(u,v)$, through which one can construct a corresponding representation in terms of the Brownian bridge sequence $\{B_n\}$, used in the representation of the quantile process. Secondly, we are aiming at those statistics $T_n(u,v)$, which could be used in estimating the location and scale parameters μ,σ of the family F of (1.1).

Now, let $J(.)$ be a real valued function defined on $[0,1]$. In the following we list a set of conditions which will be used in the present section. We emphasize that only subsets of it will be used at different stages in the sequel.

$$\left.\begin{array}{l}
\text{(i)} \quad \int J(x)\, d\, G^{-1}(x) < \infty \\[2ex]
\text{(ii)} \quad \sup_{0 \le u < v \le 1} \left| \frac{1}{n} \sum_{[nu]+1}^{[nv]} J\left(\frac{j}{n+1}\right) G^{-1}\left(\frac{j}{n+1}\right) \right. \\[2ex]
\qquad\qquad \left. - \int_u^v J(x)\, G^{-1}(x)\, dx \right| = o(n^{-1/2}) \\[2ex]
\text{(iii)} \quad \text{If } \{B(x)\ ,\ 0 \le x \le 1\} \text{ is a Brownian bridge, then}
\end{array}\right\} \qquad (3.4)$$

$$\sup_{0 \le u < v \le 1} \left| \int_u^v J(x)\, B(x)\, d\, G^{-1}(x) \right| < \infty \qquad \text{a.s.}$$

As to our main result of this section, we have

<u>Theorem 3.1</u>

Let $X_1, X_2, \ldots$ be i.i.d.rv with a continuous distribution function G which is also twice differentiable on (a,b), where $-\infty \le a = \sup\{x : G(x) = 0\}$, $+\infty \ge b = \inf\{x : G(x) = 1\}$ and $g = G' \ne 0$ on (a,b). Let $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ be the order statistics of $X_1, X_2, \ldots, X_n$. Let $J(.)$ be a real valued function for which conditions (3.4) (i) and (ii) holds true. One can then define a Brownian brdige $\{B_n(x)\ ;\ 0 \le x \le 1\}$ for each n such that if condition (1.6) of Theorem 1A is assumed then

$$\sup_{\delta_n \le u < v \le 1-\delta_n} \left| n^{1/2} \left(T_n(u,v) - \int_u^v J(x)\, G^{-1}(x)\, dx \right) \right.$$

$$\left. - \frac{1}{n} \sum_{[nu]+1}^{[nv]} \frac{J(\frac{j}{n+1})\, B_n(\frac{j}{n+1})}{g(G^{-1}(\frac{j}{n+1}))} \right| \underline{\text{a.s.}}\ o(1), \quad (3.5)$$

where $\delta_n = 25\, n^{-1} \log \log n$.

If, in addition to (1.6), condition (1.9) of Theorem 1A is also assumed, then

$$\sup_{0 \le u < v \le 1} \left| n^{1/2} \left(T_n(u,v) - \int_u^v J(x)\, G^{-1}(x)\, dx \right) \right.$$

$$\left. - \frac{1}{n} \sum_{[nu]+1}^{[nv]} \frac{J(\frac{j}{n+1})\, B_n(\frac{j}{n+1})}{g(G^{-1}(\frac{j}{n+1}))} \right| \underline{\text{a.s.}}\ o(1). \quad (3.6)$$

<u>Proof</u>

First, we prove (3.5). Since

$$T_n(u,v) = n^{-1} \sum_{[nu]+1}^{[nv]} J(\frac{j}{n+1}) X_{j:n}$$

$$= n^{-1} \sum_{[nu]+1}^{[nv]} J\left(\frac{j}{n+1}\right) \frac{n^{1/2} g\left(G^{-1}\left(\frac{j}{n+1}\right)\right)\left(X_{j:n} - G^{-1}\left(\frac{j}{n+1}\right)\right)}{n^{1/2} g\left(G^{-1}\left(\frac{j}{n+1}\right)\right)}$$

$$+ n^{-1} \sum_{[nu]+1}^{[nv]} J\left(\frac{j}{n+1}\right) \cdot G^{-1}\left(\frac{j}{n+1}\right) .$$

Then

$$n^{1/2} \left[T_n(u,v) - \int_u^v J(x)\, G^{-1}(x)\, dx\right]$$

$$= n^{-1} \sum_{j=[nu]+1}^{[nv]} J\left(\frac{j}{n+1}\right) \rho_n\left(\frac{j}{n+1}\right) / g\left(G^{-1}\left(\frac{j}{n+1}\right)\right)$$

$$+ n^{1/2} \left[n^{-1} \sum_{[nu]+1}^{[nv]} J\left(\frac{j}{n+1}\right) G^{-1}\left(\frac{j}{n+1}\right)\right.$$

$$\left. - \int_u^v J(x)\, G^{-1}(x)\, dx\right] . \tag{3.7}$$

As to the first summation on the right hand side of (3.7), we have by (1.7) of Theorem 1A, for any u,v s.t $\delta_n \le u < v \le 1-\delta_n$

$$n^{-1} \sum_{[nu]+1}^{[nv]} J\left(\frac{j}{n+1}\right) \rho_n\left(\frac{j}{n+1}\right) / g\left(G^{-1}\left(\frac{j}{n+1}\right)\right)$$

$$\underline{\text{a.s.}} \quad n^{-1} \sum_{[nu]+1}^{[nv]} J\left(\frac{j}{n+1}\right) B_n\left(\frac{j}{n+1}\right) / g\left(G^{-1}\left(\frac{j}{n+1}\right)\right)$$

$$+ O(n^{-1/2} \log n)\, n^{-1} \sum_{[nu]+1}^{[nv]} J\left(\frac{j}{n+1}\right) / g\left(G^{-1}\left(\frac{j}{n+1}\right)\right)$$

$$\underline{\text{a.s.}} \quad n^{-1} \sum_{[nu]+1}^{[nv]} J\left(\frac{j}{n+1}\right) B_n\left(\frac{j}{n+1}\right) / g\left(G^{-1}\left(\frac{j}{n+1}\right)\right)$$

$$+ O(n^{-1/2} \log n) \left\{ \int_u^v J(x)\, d\, G^{-1}(x) + o(n^{-1/2}) \right\} . \tag{3.8}$$

So that, by (3.4) (i), the right hand side of (3.8) is

$$n^{-1} \sum_{[nu]+1}^{[nv]} J\left(\frac{j}{n+1}\right) B_n\left(\frac{j}{n+1}\right) / g\left(G^{-1}\left(\frac{j}{n+1}\right)\right) + O(n^{-1/2} \log n).$$

As to the second term on the right hand side of (3.7), we have by (3.4) (ii)

$$n^{1/2}\left[n^{-1}\sum_{[nu]+1}^{[nv]} J\left(\frac{j}{n+1}\right) G^{-1}\left(\frac{j}{n+1}\right)\right.$$

$$\left. - \int_u^v J(x)\, G^{-1}(x)\, dx \right] = o(1) .$$

Hence

$$\sup_{\delta_n \le u < v \le 1-\delta_n} \left| n^{1/2} \left[T_n(u,v) - \int_v^u J(x)\, G^{-1}(x)\, dx \right] \right.$$

$$\left. - \frac{1}{n} \sum_{[nu]+1}^{[nv]} J\left(\frac{j}{n+1}\right) B_n\left(\frac{j}{n+1}\right) / g\left(G^{-1}\left(\frac{j}{n+1}\right)\right) \right| \underline{\underline{a.s.}}\ o(1) .$$

Similarly, we can prove (3.6).

<u>Corollary 3.1</u>

Under all the conditions of Theorem 3.1, but condition (1.9) of Theorem 1A, we have

$$\sup_{\delta_n \le u < v \le 1-\delta_n} \left| n^{1/2} \left[ET_n(u,v) - \int_v^u J(x)\, G^{-1}(x)\, dx \right] \right| = o(1) \quad (3.9)$$

If, condition (1.9) of Theorem 1A is also assumed, then

$$\sup_{0 \le u < v \le 1} \left| n^{1/2} \left[ET_n(u,v) - \int_u^v J(x)\, G^{-1}(x)\, dx \right] \right| = o(1) . \quad (3.10)$$

Corollary 3.2

Assume that condition (3.4) (iii) is satisfied. Then under all the conditions of Theorem 1A, but condition (1.9), we have

$$\sup_{\delta_n \le u < v \le 1-\delta_n} \left| n^{1/2} \left(T_n(u,v) - \int_u^v J(x)\, G^{-1}(x)\, dx \right) \right.$$

$$\left. - \int_u^v J(x)\, B_n(x)\, d\, G^{-1}(x) \right| \overset{P}{=} o(1) . \tag{3.11}$$

In addition, if condition (1.9) of Theorem 1A is also satisfied. Then

$$\sup_{0 \le u \le v \le 1} \left| n^{1/2} \left(T_n(u,v) - \int_u^v J(x)\, G^{-1}(x)\, dx \right) \right.$$

$$\left. - \int_u^v J(x)\, B_n(x)\, d\, G^{-1}(x) \right| \overset{P}{=} o(1) . \tag{3.12}$$

Remark 3.1

If condition (3.4) (iii) is not satisfied, we can replace it by

$$\sup_{\delta \le u < v \le 1-\delta} \left| \int_u^v J(x)\, B(x)\, d\, G^{-1}(x) \right| < \infty \quad \text{a.s.,} \tag{3.4)(iii)*}$$

where $\delta \in (0,1/2)$. Then there exists n_0 such that $\delta_{n_0} \le \delta$, and

hence, we have

$$\sup_{\delta \le u < v \le 1-\delta} \left| n^{1/2} \left(T_n(u,v) - \int_u^v J(x)\, G^{-1}(x)\, dx \right) \right.$$

$$\left. - \int_u^v J(x)\, B_n(x)\, d\, G^{-1}(x) \right| \overset{P}{=} o(1) . \tag{3.13}$$

<u>Remark 3.2</u>

A sufficient condition for (3.4) (iii) to be satisfied is

$$\int_0^1 \left\{ x(1-x) \log\log \frac{1}{x(1-x)} \right\}^{1/2} |J(x)|\, d\, G^{-1}(x) < \infty , \tag{3.14}$$

for,

$$\sup_{0 \le u < v \le 1} \left| \int_u^v J(x)\, B(x)\, d\, G^{-1}(x) \right|$$

$$\le \sup_{0<y<1} \frac{|B(y)|}{\{y(1-y)\log\log \frac{1}{y(1-y)}\}^{1/2}} \cdot$$

$$\int_0^1 \left\{ x(1-x) \log\log \frac{1}{x(1-x)} \right\}^{1/2} |J(x)|\, d\, G^{-1}(x) < \infty ,$$

by (3.14), and since by the law of iterated logarithm and the almost sure continuity of $B(y)$,

$$\sup_{0<y<1} \frac{|B(y)|}{\{y(1-y)\log\log\frac{1}{y(1-y)}\}^{1/2}}$$

is almost surely finite.

Remark 3.3

The conditions of Theorem (3.1) including condition (3.14) are more or less similar to those conditions used by Chernoff, Castwirth, Johns (1967) to prove (3.2). Indeed, they assumed (3.4) (i) and (ii), $\sigma^2 < +\infty$ (σ^2 is as in (3.3)), tail smoothness conditions, and that

$$\left| \int_0^1 J(u)\ \{u(1-u)\}^{1/2}\ dG^{-1}(u) \right| < +\infty \quad . \tag{3.15}$$

On the other hand Theorem (3.1) requires all the above mentioned conditions except that $\sigma^2 < +\infty$ and (3.15) are replaced by (3.14). This results from the introduction of the Brownian bridge in the representations (3.11) and (3.12) and from Remark (3.2).

Example 3.1 The Trimmed Mean.

For any $0 < u < \frac{1}{2}$ define

$$T_n(u) = \frac{1}{n - 2[nu]} \sum_{[nu]+1}^{n-[nu]} X_{j:n}$$

$$= \frac{n}{n - 2[nu]} \left\{ \frac{1}{n} \sum_{[nu]+1}^{n-[nu]} X_{j:n} \right\}$$

$$= \frac{1}{1 - 2\frac{[nu]}{n}} \left\{ \frac{1}{n} \sum_{[nu]+1}^{n-[nu]} X_{j:n} \right\} . \tag{3.16}$$

By Theorem (3.1), we have

$$n^{1/2} \left[T_n(u) - \frac{1}{1 - 2u} \int_u^{1-u} G^{-1}(x)\, dx \right]$$

$$\xrightarrow{D} \frac{1}{1 - 2u} \int_u^{1-u} \frac{B(x)}{g(G^{-1}(x))}\, dx$$

$$= \frac{1}{1 - 2u} \int_u^{1-u} B(x) dG^{-1}(x) . \tag{3.17}$$

For the sake of future reference , we recall here the standard conditions for the validity of the Cramér-Rao bounds as:

(i) The parameter space Θ is open in R^2

(ii) $f(x;\theta)$ is positive on a set S independent of $\theta \in \Theta$.

(iii) $(\frac{\partial}{\partial\theta})(f(x;\theta)$ exists for all $\theta \in \Theta$ and $x \in S$ with probability one.

(iv) $\int_S \cdots \int_S f(x_i;\theta) \ldots f(x_n;\theta)\, dx_1 \ldots dx_n$ may be differentiated under the integral sign. (3.18)

(v) Fisher information ~~matrix I^*~~ (cf. (3.19)) is non-singular.

(vi) $\int_S \cdots \int_S U(x_1, \ldots, x_n)\, f(x_1;\theta) \ldots f(x_n;\theta) dx_1 \ldots dx_n$

may be differentiated under the integral sign, where $U(X_1, X_2, \ldots, X_n)$ is any unbiased estimator of $\xi(\theta)$.

Now, we consider the case where the sample $X_1, X_2, \ldots, X_n$ is taken from $F \in \mathcal{F}$ of (1.1) with F_0 specified and satisfies all the standard conditions for the validity of the Cramér-Rao bound i.e., conditions (3.18) (i)-(vi). In the following we apply Theorem (3.1) to Bennett's (1952) estimates of location and scale parameters, whose asymptotic efficiency is well known (cf. Chernoff, Gastwirth, Johns (1967)). Our aim here is to study the asymptotic representation of these estimates in terms of Theorem (3.1). Towards this end, we recall that Fisher information matrix I^* in this case is given

by $I^* = \sigma^{-2} I$, where

$$I = \begin{bmatrix} \int_{-\infty}^{\infty} L_1'(y) \; f_0(y)dy & \int_{-\infty}^{\infty} L_2'(y) \; f_0(y)dy \\ \int_{-\infty}^{\infty} yL_1'(y) \; f_0(y)dy & \int_{-\infty}^{\infty} yL_2'(y) \; f_0(y)dy \end{bmatrix} ,$$

with

$$L_1(y) = -f_0'(y)/f_0(y) \tag{3.19}$$

$$L_2(y) = yL_1(y) - 1$$

In addition, we recall that Bennett's (1952) estimates of the location and scale parameters are as follows:

<u>Location parameter</u> (scale parameter σ known)

$$\hat{\mu}_{n1} = T_{n11} - \sigma I_{12}/I_{11} \quad , \tag{3.20}$$

where

$$T_{n11} = \frac{1}{n} \sum_{j=1}^{n} J_{11}\left(\frac{j}{n+1}\right) X_{j:n} \quad , \tag{3.21}$$

and

$$J_{11}(u) = I_{11}^{-1} L_1'(y) \quad , \quad y = F_0^{-1}(u) \quad . \tag{3.22}$$

<u>Scale parameter</u> (location parameter μ known)

$$\hat{\sigma}_{n1} = T_{n12} - \mu I_{12}/I_{22} \quad , \tag{3.23}$$

where

$$T_{n12} = \frac{1}{n} \sum_{j=1}^{n} J_{12}\left(\frac{j}{n+1}\right) X_{j:n} \quad , \tag{3.24}$$

and

$$J_{12}(u) = I_{22}^{-1} L_2'(y) \quad , \; y = F_0^{-1}(u) \quad . \tag{3.25}$$

<u>Both parameters unknown</u>

$$(\hat{\mu}_{n2}, \hat{\sigma}_{n2}) = (T_{n21}, T_{n22}) \quad , \tag{3.26}$$

where

$$\left.\begin{aligned} T_{n2i} &= \frac{1}{n} \sum_{j=1}^{n} J_{2i}\left(\frac{j}{n+1}\right) X_{j:n} \quad , \; i = 1,2 \\ J_{21}(u) &= \{I_{11}I_{22}J_{11}(u) - I_{12}I_{22}J_{12}(u)\}/(I_{11}I_{22} - I_{12}^2) \; , \\ J_{22}(u) &= \{I_{11}I_{22}J_{12}(u) - I_{12}I_{11}J_{11}(u)\}/(I_{11}I_{22} - I_{12}^2) \; . \end{aligned}\right\} \tag{3.27}$$

Remark 3.4

If F_0 is symmetric, then $I_{12} = 0$ and consequently $\hat{\mu}_{n1} = \hat{\mu}_{n2}$ i.e., the efficient estimate of μ is the same regardless of whether or not σ is known. The same is true for the efficient estimate of σ i.e., $\hat{\sigma}_{n1} = \hat{\sigma}_{n2}$.

We also have the following Lemma,

Lemma 3.1

Let $X_1, X_2, \ldots$ be i.i.d.rv with a continuous distribution function $F \in \mathcal{F}$ of (1.1). Suppose that F_0 satisfies all the conditions of Theorem (3.1) as well as the standard conditions for the validity of the Cramér-Rao bounds (cf. (3.18)). Then one can define a Brownian bridge $\{B_n(x) ; 0 \leq x \leq 1\}$ for each n such that

$$n^{1/2}(\hat{\mu}_{ni} - \mu)/\sigma = \int_0^1 J_{i1}(x) B_n(x)\, dF_0^{-1}(x) + o_p(1) , \quad i=1,2 , \tag{3.28}$$

and

$$n^{1/2}\left(\frac{\sigma_{ni} - \sigma}{\sigma}\right) = \int_0^1 J_{i2}(x) B_n(x)\, dF_0^{-1}(x) + o_p(1) , \quad i=1,2 . \tag{3.29}$$

Remark 3.5

It should be noticed that the sequence $\{B_n(x) ; 0 \leq x \leq 1\}$ of Brownian bridges in Lemma (3.1) is the same sequence that has been

used to approximate the standardized quantile process.

4. The Asymptotic Distribution of $M_n(\lambda)$ when using L-Estimators

Let $X_1, X_2, \ldots$ be a sequence of i.i.d.rv's with a continuous distribution function $F \in \mathcal{F}$ of (1.1) with F_0 specified. Suppose that F_0 satisfies all the conditions of Lemma (3.1). In the following we will consider the general case where both μ and σ are unknown. Let $(\hat{\mu}_{n2}, \hat{\sigma}_{n2})$ be as in (3.26), and define

$$T_n^{(i1)} = \int_0^1 J_{i1}(x)\, B_n(x)\, d\, F_0^{-1}(x) \qquad i = 1,2 \quad , \tag{4.1}$$

and

$$T_n^{(i2)} = \int_0^1 J_{i2}(x)\, B_n(x)\, d\, F_0^{-1}(x) \qquad i = 1,2 \quad , \tag{4.2}$$

where J_{i1}, J_{i2}, $i = 1, 2$ are as defined in (3.22), (3.25) and (3.27). And $\{B_n(x) \; ; \; 0 \le x \le 1\}$ is that sequence of Brownian bridges of Theorem 1A.

Lemma 4.1

Under all the conditions of Lemma (3.1) the estimated standardized empirical quantile process $\hat{\rho}_n^0(y)$ (cf. (1.10)), can be written in terms of the standardized empirical quantile process $\rho_n^0(y)$ as follows:

$$\hat{\rho}_n(y) = \frac{1}{1 + o_p(1)} \{\rho_n^0(y) - T_n^{(21)} f_0(F_0^{-1}(y))$$

$$- T_n^{(22)} f_0(F_0^{-1}(y))\, F_0^{-1}(y) + o_p(1)\, (f_0(F_0^{-1}(y))$$

$$+ f_0(F_0^{-1}(y)) F_0^{-1}(y))\}. \tag{4.3}$$

Proof

Similar to that of Corollary (2.1), except that we use Lemma (3.1) in place of Lemma (2.1).

The main result of this section is

Theorem 4.1

Let $X_1, X_2, \ldots, X_n$ be a random sample with a distribution function $F \in \mathcal{F}$ of (1.1). Let $(\hat{\mu}_{n2} \quad \hat{\sigma}_{n2})$ be the sequence of estimators of $(\mu \quad \sigma)$ based on $X_1, X_2, \ldots, X_n$, which is as defined in (3.26). Assume also that F_0 satisfies all the conditions of Lemma (3.1) including conditions (1.9) of Theorem 1A. Further assume that

$$\int_{-\infty}^{\infty} f_0^3(x)\, dx < +\infty$$

Then, there exists a sequence of Brownian bridges $\{B_n\}$ such that for $\lambda = 1, 2, \ldots$, we have

$$|M_n^{(2)}(\lambda) - G_n^{(2)}(\lambda)| \xrightarrow{P} 0 \quad , \tag{4.4}$$

where

$$G_n^{(2)}(\lambda) = \int_0^1 B_n^2(y)\lambda^{-1}\, d(F_0^{-1}(y))^{\lambda} + (T_n^{(21)})^2\, J^{(\lambda-1)}$$

$$+ (T_n^{(22)})^2\, J^{(\lambda+1)} - 2(T_n^{(21)}\, R_n^{(\lambda-1)}$$

$$+ T_n^{(22)}\, R_n^{(\lambda)} - T_n^{(21)} T_n^{(22)}\, J^{(\lambda)}) \quad , \tag{4.5}$$

where $J^{(\alpha)}$, $\alpha = \lambda-1, \lambda, \lambda+1$ is as in (2.15), $R_n^{(\alpha)}$, $\alpha = \lambda-1, \lambda$ is as in (2.16) and $T_n^{(21)}$, $T_n^{(22)}$ is as in (4.1), (4.2) respectively.

The proof is similar to that of Theorem (2.1).

<u>Corollary 4.1</u>

The conditions of Theorem (4.1) imply

$$M_n^{(2)}(\lambda) \xrightarrow{D} G^{(2)}(\lambda) \quad , \quad \lambda = 1,2,.. \quad , \tag{4.5}$$

where

$$G^{(2)}(\lambda) = \int_0^1 B^2(y)\lambda^{-1}\, d((F_0^{-1}(y))^{\lambda}) + (T^{(21)})^2\, J^{(\lambda-1)}$$

$$+(T^{(22)})^2\, J^{(\lambda+1)} - 2(T^{(21)}\, R^{(\lambda-1)} + T^{(22)}\, R^{(\lambda)}$$

$$- T^{(21)}\, T^{(22)}\, J^{(\lambda)}) \quad , \tag{4.6}$$

where

$$T^{(2j)} = \int_0^1 J_{2j}(x)\, B(x)\, d\, F_0^{-1}(x) \quad , \quad j = 1,2 \quad , \tag{4.7}$$

and

$$R^{(\alpha)} = \int_0^1 B(y)\ (F_0^{-1}(y))^{\alpha}\ dy \qquad , \quad \alpha = \lambda-1,\ \lambda \quad . \tag{4.8}$$

<u>Corollary 4.2</u>

In the case where μ is unknown and σ is known, we have

$$M_n^{(2)}(\lambda) \xrightarrow{D} \int_0^1 B^2(y)\lambda^{-1}\ d(F_0^{-1}(y))^{\lambda} + (T^{(11)})^2\ J^{(\lambda+1)}$$

$$- 2T^{(11)}\ R^{(\lambda-1)} \qquad , \tag{4.9}$$

while, if μ is assumed to be known, then we have

$$M_n^{(2)}(\lambda) \xrightarrow{D} \int_0^1 B^2(y)\lambda^{-1}\ d(F_0^{-1}(y))^{\lambda} + (T^{(12)})^2\ J^{(\lambda+1)}$$

$$- 2T^{(12)}\ R^{(\lambda)} \qquad , \tag{4.10}$$

where

$$T^{(1j)} = \int_0^1 J_{1j}(x)\ B(x)\ d\ F_0^{-1}(x) \qquad , \quad j = 1,2 \quad .$$

The Asymptotic Distribution of $M_n(\lambda)$ when using Weiss' Estimate

Let $X_1, X_2, \ldots$ be a sequence of i.i.d.r.v's with a continuous distribution function $F \in \mathcal{F}$ of (1.1). In this section we will consider the case where μ is known, and without loss of generality we may take $\mu = 0$. Weiss (1961, 1963) suggested the following estimate of the scale parameter

$$\hat{\sigma}_n = \sum_{j=1}^{n-1} f_0(F_0^{-1}(\frac{j}{n+1}))\,(X_{j+1:n} - X_{j:n}) \quad . \tag{5.1}$$

In addition, Weiss (1961) proved that $\hat{\sigma}_n$ of (5.1) is a consistent estimate of σ. Csörgő, Révész (1980 I,II) have proved the following

Lemma 2B(Csörgő, Révész (1980 I,II)).

Assume that for F_0 of the family $\mathcal{F}$ of (1.1) conditions (1.6) and (1.9) of Theorem 1A hold true, and that

$$f_0' \text{ is continuous on } (a,b) \quad , \tag{5.2}$$

then

$$n^{1/2}(\hat{\sigma}_n - \sigma)/\sigma = -\int_0^1 \frac{f_0'(F_0^{-1}(y))}{f_0^2(F_0^{-1}(y))}\, B_n(y)\, dy + o_p(1)$$

$$= T_n^{(3)} + o_p(1) \quad , \tag{5.3}$$

where $\{B_n\}$ is the same sequence of Brownian bridges of Theorem 1A.

Corollary 5.1

Under the conditions of Lemma 2B, the estimated standardized empirical quantile process $\hat{\rho}_n^0(y)$ of (1.10), can be written in terms of the standardized empirical quantile process as follows:

$$\hat{\rho}_n^0(y) = \frac{1}{1 + o_p(1)} \left\{ \rho_n^0(y) - T_n^{(3)} f_0(F_0^{-1}(y)) F_0^{-1}(y) + o_p(1) f_0(F_0^{-1}(y)) F_0^{-1}(y) \right\} . \tag{5.3}$$

Proof

Similar to that of corollary (2.1), except that Lemma 2B is used in place of Lemma (2.1).

In the following theorem, we give the asymptotic distribution of $M_n(\lambda)$ when using Weiss' estimate of (5.1).

Theorem 5.1

Assume that for F_0 of the family F of (1.1) conditions (1.6) and (1.9) of Theorem 1A hold true, and that condition (5.2) of Lemma 2B holds true. Further, assume that condition (2.11) of theorem (2.1) holds true, then there exists a sequence of Brownian bridges $\{B_n\}$ such that for $\lambda = 1, 2, \ldots$, we have

$$|M_n^{(3)}(\lambda) - G_n^{(3)}(\lambda)| \xrightarrow{P} 0 \quad , \tag{5.4}$$

where

$$G_n^{(3)}(\lambda) = \int_0^1 B_n^2(y)\lambda^{-1}\, d(F_0^{-1}(y))^\lambda$$

$$+ (T_n^{(3)})^2\, J^{(\lambda+1)} - 2(T_n^{(3)}\, R_n^{(\lambda)}) \quad ,$$

where $T_n^{(3)}$ is as defined in (5.3), and $J^{(\lambda+1)}$ and $R_n^{(\lambda)}$ are as in (2.14) and (2.15) respectively.

<u>Corollary 5.2</u>

The conditions of Theorem (5.1) imply

$$M_n^{(3)}(\lambda) \xrightarrow{D} G^{(3)}(\lambda) \quad , \quad \lambda = 1, 2, \ldots \quad , \tag{5.5}$$

where

$$G^{(3)}(\lambda) = \int_0^1 B^2(y)\lambda^{-1}\, d(F_0^{-1}(y))^\lambda$$

$$+ (T^{(3)})^2\, J^{(\lambda+1)} - 2T^{(3)}\, R^{(\lambda)} \quad ,$$

where

$$T^{(3)} = -\int_0^1 \frac{f_0'(F_0^{-1}(y))}{f_0^2(F_0^{-1}(y))}\, B(y)dy \quad , \tag{5.6}$$

and

$$R^{(\lambda)} = \int_0^1 B(y)\, (F_0^{-1}(y))^\lambda dy \quad .$$

6. Discussion and Examples

It is clear, from sections 2, 4 and 5 together with Theorem (3.1) of Csörgő, Révész (1979), that for any "reasonable" method of estimation, like for example $(\hat{\mu}_n \quad \hat{\sigma}_n) = (\bar{X}_n \quad S_n)$, maximum likelihood and maximum likelihood-like estimators, L-estimators, or the Weiss' estimate for σ when μ is known, we have

$$M_n(\lambda) \xrightarrow{D} G(\lambda) \quad , \quad \lambda = 1,2,\ldots \quad , \tag{6.1}$$

where the specific forms of $G(\lambda)$ and $M_n(\lambda)$ are to be determined according to the method of estimation. But in all cases, the distribution of $G(\lambda)$ is nuisance parameter free and depends only on the given form of F_0 in (1.2). This means that via (6.1) we have a possibility of testing for H_0 of (1.2), provided that we can evaluate the distribution of $G(\lambda)$ for a given F_0. However this task will not be simple in general. The distribution of $G(\lambda)$ is somewhat simpler when F_0 is symmetric around zero. In the latter case $J^{(\lambda)} = 0$ if λ is an odd integer, and $J^{(\lambda-1)} = J^{(\lambda+1)} = 0$ if λ is an even integer, where $J^{(\alpha)}$, $\alpha = \lambda-1, \lambda, \lambda+1$ are as in (2.14).

In discussing the consistency of $M_n(\lambda)$, we observe that $M_n(\lambda)$ is a real valued random variable when λ is an even integer and it is a positive random variable when λ is an odd integer. Hence if λ is an odd integer, we should reject H_0 of (1.2) when $M_n(\lambda)$ is too large, if λ is an even integer, we should reject H_0 of (1.2) when $|M_n(\lambda)|$ is too large. In addition, if λ is an odd integer and if H_0 is not true then $M_n(\lambda) \xrightarrow{a.s.} \infty$. The latter property is not true when λ is

an even integer (consider F_0 to be symmetric around zero and take any symmetric alternative). To sum up: if λ is an odd integer then tests based on $M_n(\lambda)$ are consistent against any alternative hypothesis, whereas, if λ is an even integer then tests based on $M_n(\lambda)$ are not necessarily consistent against all alternatives.

Example 1: On Testing for Exponentiality

Consider the family of exponential density functions

$$EXP(\mu,\sigma) = \left\{ f(x;\mu,\sigma) : f(x;\mu,\sigma) = f_0\left(\frac{x-\mu}{\sigma}\right) \right.$$

$$\left. = \frac{1}{\sigma} \exp\left(-\frac{x-\mu}{\sigma}\right);\ x \geq \mu \ \varepsilon\ R^1,\ \sigma > 0 \right\} . \tag{6.2}$$

Let $X_1, X_2, \ldots, X_n$ $(n \geq 2)$ be a random sample with a density function f, we wish to test the following composite null hypotheses

$$H_0^{(1)} : f \ \varepsilon\ EXP(0,\sigma) \quad , \quad \sigma > 0 \quad , \tag{6.3}$$

and

$$H_0^{(2)} : f \ \varepsilon\ EXP(\mu,\sigma) \quad , \quad \mu \ \varepsilon\ R^1 \quad , \sigma > 0 \quad . \tag{6.4}$$

First we consider $H_0^{(1)}$ of (6.3). Under $H_0^{(1)}$, $X_1, X_2, \ldots$ $\ldots, X_n$ are independent positive random variables with density function $f \ \varepsilon\ EXP(0,\sigma)$, $\sigma > 0$. In the following, we consider three possible cases.

Case 1

Using the maximum likelihood estimator of σ of $EXP(0,\sigma)$; i.e., take $\hat{\sigma}_n^{(1)} = \bar{X}_n$. Consequently

$$M_n^{(1)}(\lambda) = \sum_{k=1}^{n} \left\{ \frac{X_{k:n}}{\bar{X}_n} - \log \frac{1}{1 - \frac{k}{n+1}} \right\}^2 \cdot$$

$$\left(1 - \frac{k}{n+1}\right) \left[\log \frac{1}{1 - \frac{k}{n+1}}\right]^{\lambda-1} . \tag{6.5}$$

It is clear that all the conditions of Theorem (2.1) are satisfied, and hence we have

$$|M_n^{(1)}(\lambda) - G_n^{(1)}(\lambda)| \xrightarrow{P} 0 \quad \text{as} \quad n \to \infty \ ; \ \lambda = 1,2,\ldots , \tag{6.6}$$

where

$$G_n^{(1)}(\lambda) = \int_0^1 B_n^2(y)\lambda^{-1} \, d(F_0^{-1}(y))^{\lambda}$$

$$+ (T_n^{(2)})^2 \, J^{(\lambda+1)} - 2 \, T_n^{(2)} \, R_n^{(\lambda)} , \tag{6.7}$$

with

$$T_n^{(2)} = \int_0^1 (F_0^{-1}(x) - 1) \, d_x \, B_n(x)$$

$$= - \int_0^{\infty} B_n(x) \, d \, F_0^{-1}(x) = - \int_0^1 \frac{B_n(x)}{f_0(F_0^{-1}(x))} \, dx , \tag{6.8}$$

where the second equality above is obtained by integrating by parts and using condition (2.1)(v). In addition

$$G_n^{(1)}(\lambda) \overset{D}{=} G^{(1)}(\lambda) = \int_0^1 \frac{B^2(y)}{1-y}\,dy + \left(\int_0^1 \frac{B(y)}{1-y}\,dy\right)^2 \cdot$$

$$\int_0^1 (1-y)(\log\frac{1}{1-y})^{\lambda+1}\,dy + 2\int_0^1 \frac{B(y)}{1-y}\,dy \;\cdot$$

$$\int_0^1 B(y)(\log\frac{1}{1-y})^{\lambda}\,dy \qquad , \lambda = 1,2,\ldots \qquad (6.9)$$

Hence

$$M_n^{(1)}(\lambda) \overset{D}{\to} G^{(1)}(\lambda) \qquad (6.10)$$

Case 2

Using the L-estimator of σ of $EXP(0,\sigma)$ which for the family $EXP(0,\sigma)$ is equal to $\bar{X}_n$, the M.L.E. So that in this case $M_n^{(2)}(\lambda) = M_n^{(1)}(\lambda)$ of (6.4). By Theorem (4.1)

$$|M_n^{(2)}(\lambda) - G_n^{(2)}(\lambda)| \overset{P}{\to} 0 \qquad , \qquad (6.11)$$

where

$$G_n^{(2)}(\lambda) = \int_0^1 B_n^2(y)\lambda^{-1}\, d(F_0^{-1}(y))^{\lambda}$$

$$+ (T_n^{(12)})^2 J^{(\lambda+1)} - 2\, T_n^{(12)} R_n^{(\lambda)} , \tag{6.12}$$

with

$$T_n^{(12)} = \int_0^1 \frac{B_n(x)}{1-x}\, dx \tag{6.13}$$

and hence

$$M_n^{(2)}(\lambda) \xrightarrow{D} G^{(2)}(\lambda) , \tag{6.14}$$

where

$$G^{(2)}(\lambda) = \int_0^1 \frac{B^2(y)}{1-y} \left(\log \frac{1}{1-y}\right)^{\lambda-1} dy$$

$$+ \left(\int_0^1 \frac{B(y)}{1-y}\, dy \right)^2 \left(\int_0^1 (1-y)\left(\log \frac{1}{1-y}\right)^{\lambda+1} dy \right)$$

$$- 2 \int_0^1 \frac{B(y)}{1-y}\, dy \int_0^1 B(y)\left(\log \frac{1}{1-y}\right)^{\lambda} dy ;$$

$$\lambda = 1, 2, \ldots \tag{6.15}$$

Case 3

Using Weiss' estimate of σ , namely

$$\hat{\sigma}_n^{(3)} = \sum_{j=1}^{n-1} f_0(F_0^{-1}(\frac{j}{n+1}))\ (X_{j+1:n} - X_{j:n})$$

$$= \frac{1}{n+1}\ (n\ \bar{X}_n + X_{n:n} - (n+1)X_{1:n}) \quad , \tag{6.16}$$

which is essentially the M.L.E. of σ .

Now

$$M_n^{(3)}(\lambda) = \sum_{k=1}^{n} \left\{ \frac{X_{k:n}}{\hat{\sigma}_n^{(3)}} - \log \frac{1}{1 - \frac{k}{n+1}} \right\}^2$$

$$(1 - \frac{k}{n+1})(\log \frac{1}{1 - \frac{k}{n+1}})^{\lambda} \quad . \tag{6.17}$$

By Theorem (5.1),

$$M_n^{(3)}(\lambda) \xrightarrow{D} G^{(3)}(\lambda) \quad , \tag{6.18}$$

where

$$G^{(3)}(\lambda) = \int_0^1 \frac{B^2(y)}{1 - y}\ (\log \frac{1}{1 - y})^{\lambda - 1}\ dy$$

$$+ (T^{(3)})^2\ J^{(\lambda+1)} - 2\ T^{(3)}\ R^{(\lambda)} \quad , \tag{6.19}$$

with

$$T^{(3)} = \int_0^1 \frac{B(y)}{1-y}\,dy \quad .$$

From the above considerations, it is clear that, when testing for $H_0^{(1)}$ of (6.2), $M_n^{(1)}(\lambda) = M_n^{(2)}(\lambda)$ and $M_n^{(3)}(\lambda)$ is essentially equal to $M_n^{(1)}(\lambda) = M_n^{(2)}(\lambda)$. In addition

$$G_n^{(1)}(\lambda) \overset{D}{=} G_n^{(2)}(\lambda) \overset{D}{=} G_n^{(3)}(\lambda) \quad .$$

For these reasons, it appears that in this case the M.L.E. version of $M_n(\lambda)$ i.e., $M_n^{(1)}(\lambda)$ has an advantage over $M_n^{(3)}(\lambda)$, namely it is simpler to calculate.

In testing for $H_0^{(2)}$ of (6.3), one can make the following transformation: $X_1-X_{1:n}, X_2-X_{1:n}, \ldots, X_n-X_{1:n}$ and delete from it the term which equals to zero (with probability one there is one and only one such term). Let us denote the new variables by $X_1^*, X_2^*, \ldots$ $\ldots, X_{n-1}^*$ and let $X_{1:n}^*, X_{2:n}^*, \ldots, X_{n-1:n}^*$ be their order statistics. Given $H_0^{(2)}$ of (6.13), $X_1^*, X_2^*, \ldots, X_{n-1}^*$ are independent EXP(0,σ)rv's, and the above procedure thus modified, can now be used to test for $H_0^{(2)}$ of (6.3). Our test statistics in this case is

$$M_n^{(1)}(\lambda) = \sum_{k=2}^{n} \left\{ \frac{X_{k:n} - X_{1:n}}{\bar{X}_n - X_{1:n}} - \log \frac{1}{1 - \frac{k}{n+1}} \right\}^2$$

$$\left(1 - \frac{k}{n+1}\right) \left(\log \frac{1}{1 - \frac{k}{n+1}}\right)^{\lambda-1} \quad . \qquad (6.20)$$

and

$$M_n^{(1)}(\lambda) \xrightarrow{D} \int_0^1 \frac{B^2(y)}{1-y} \left(\log \frac{1}{1-y}\right)^{\lambda-1} dy$$

$$+ \left(\int_0^1 \frac{B(y)}{1-y}\, dy \right)^2 \int_0^1 (1-y)\left(\log \frac{1}{1-y}\right)^{\lambda+1} dy$$

$$- 2 \int_0^1 \frac{B(y)}{1-y}\, dy \int_0^1 B(y)\left(\log \frac{1}{1-y}\right)^{\lambda} dy$$

$$\lambda = 1,2\ldots \qquad . \qquad (6.21)$$

<u>Example 2:</u> On Testing for the Logistic Distribution

Consider the family of logistic density functions

$$\mathrm{LOGIS}(\mu,\sigma) = \{f(x;\mu,\sigma) : f(x;\mu,\sigma) = f_0\left(\frac{x-\mu}{\sigma}\right)$$

$$= \left[\exp - \left(\frac{x-\mu}{\sigma}\right)\right] / \sigma \left[1 + \exp - \left(\frac{x-\mu}{\sigma}\right)\right]^2 ;$$

$$-\infty < x < +\infty \qquad \mu \in R^1 , \quad \sigma > 0\} . \qquad (6.22)$$

Let $X_1, X_2, \ldots, X_n$ $(n \geq 2)$ be a random sample with a density function f, we wish to test the following composite null hypothesis

$$H_0 : f \in \text{LOGIS}(\mu,\sigma) \quad , \quad \mu \in R^1 , \quad \sigma > 0 \quad . \tag{6.23}$$

Under H_0 of (6.23), $X_1, X_2, \ldots, X_n$ are i.i.d.rv with density function $f \in \text{LOGIS}(\mu,\sigma)$, $\mu \in R^1$, $\sigma > 0$. It is known that the maximum likelihood estimation procedure is very complicated in this case. Therefore, we will consider two cases only.

Case 1

Using the L-estimators of μ and σ. It is clear that $f_0(x) = e^x/(1 + e^x)^2$ is symmetric and hence by Remark (3.4), we have

$$\hat{\mu}_{n1} = \hat{\mu}_{n2} = \frac{1}{n} \sum_{j=1}^{n} J_{11}\left(\frac{j}{n+1}\right) X_{j:n} \quad , \tag{6.24}$$

and

$$\hat{\sigma}_{n1} = \hat{\sigma}_{n2} = \frac{1}{n} \sum_{j=1}^{n} J_{12}\left(\frac{j}{n+1}\right) X_{j:n} \quad , \tag{6.25}$$

where

$$J_{11}(u) = 6u(1 - u) \quad , \tag{6.26}$$

and

$$J_{12}(u) = \frac{9}{3 + \Pi^2} \left\{ 2u - 1 + 2u(1 - u) \log \frac{1}{u(1 - u)} \right\} . \tag{6.27}$$

Now, it is easy to verify the required regularity assumptions of Theorem (4.1). Therefore

$$M_n^{(2)}(\lambda) \xrightarrow{D} G^{(2)}(\lambda) ,$$

where

$$G^{(2)}(\lambda) = \begin{cases} \displaystyle\int_0^1 \frac{B^2(y)}{y(1-y)} \left(\log \frac{1-y}{y}\right)^{\lambda-1} dy - 2\, T^{(22)}\, R^{(\lambda)} \\ \qquad + 2\, T^{(21)}\, T^{(22)}\, J^{(\lambda)} , \\ \qquad \text{if } \lambda \text{ is an even integer} \\ \displaystyle\int_0^1 \frac{B^2(y)}{y(1-y)} \left(\log \frac{1-y}{y}\right)^{\lambda-1} dy + (T^{(21)})^2\, J^{(\lambda-1)} \\ \qquad + (T^{(22)})^2\, J^{(\lambda+1)} \\ \qquad -2(T^{(21)} R^{(\lambda-1)} + T^{(22)} R^{(\lambda)}) , \\ \qquad \text{if } \lambda \text{ is an odd integer,} \end{cases} \tag{6.28}$$

where

$$T^{(21)} = 6 \int_0^1 B(y)dy , \tag{6.29}$$

$$T^{(22)} = \frac{9}{3 + \pi^2} \int_0^1 \left\{ \frac{2y - 1}{y(1 - y)} + 2\log \frac{1}{y(1 - y)} \right\} B(y)dy$$

$$J^{(\alpha)} = \int_{-\infty}^{\infty} y^{\alpha} \frac{e^{2y}}{(1 + e^{y})^4} dy \quad , \alpha = \lambda-1, \lambda, \lambda+1 \quad ,$$

and

$$R^{(\alpha)} = \int_0^1 B(y)(\log \frac{1 - y}{y})^{\alpha} dy \quad , \alpha = \lambda-1, \lambda \quad .$$

Case 2

Using Weiss' estimate of σ, when μ is known, i.e.,

$$\hat{\sigma}_n^{(3)} = \sum_{j=1}^{n-1} \frac{j}{n + 1}(1 - \frac{j}{n + 1})(X_{j+1:n} - X_{j:n}) \tag{6.30}$$

By Theorem (5.1), we have

$$M_n^{(3)}(\lambda) \xrightarrow{D} G^{(3)}(\lambda) \quad , \tag{6.31}$$

where

$$G^{(3)}(\lambda) \to \begin{cases} \int_0^1 \frac{B^2(y)}{y(1 - y)} (\log \frac{1 - y}{y})^{\lambda-1} dy - 2T^{(3)} R^{(\lambda)} \quad , & \text{if } \lambda \text{ is an even integer} \\ \int_0^1 \frac{B^2(y)}{y(1 - y)} (\log \frac{1 - y}{y})^{\lambda-1} dy + (T^{(3)})^2 J^{(\lambda+1)} - 2T^{(3)} R^{(\lambda)} \quad , & \text{if } \lambda \text{ is an odd integer} \end{cases}$$

where $R^{(\alpha)}$ and $J^{(\alpha)}$ are as above and

$$T^{(3)} = \int_0^1 \frac{2y - 1}{y(1 - y)} B(y)dy \; .$$

From the above considerations, when testing for H_0 of (6.23) with $\mu \in R^1$ is unknown, then we take the route of case 1 above. However, if μ is known, then it is easier to work through the argument of case 2. In both cases we should take λ to be an odd integer, since for this choice, our test procedure will be consistent against any alternative hypothesis, where as if λ is an even integer then tests based on $M_n(\lambda)$ are not consistent against symmetric alternatives.

<u>Example 3</u>: On Testing for the Weibull and Extreme Value Distributions

Consider the family of Extreme value density functions

$$\text{EXTVAL}(\mu,\sigma) = \left\{ f(x;\mu,\sigma) : f(x;\mu,\sigma) = f_0\left(\frac{x - \mu}{\sigma}\right) = \right.$$

$$\left. \frac{1}{\sigma} \exp\left(\frac{x - \mu}{\sigma}\right) \exp\left\{-\exp\left(\frac{x - \mu}{\sigma}\right)\right\} , \quad \sigma > 0, \; \mu \in R^1 \right\} . \qquad (6.32)$$

Let $X_1, X_2, \ldots, X_n$ $(n \geq 2)$ be a random sample with a density function f, we wish to test the following composite null hypothesis

$$H_0 : f \in \text{EXTVAL}(0,\sigma) \quad , \quad \sigma > 0 \; . \qquad (6.33)$$

Under H_0 of (6.33), the maximum likelihood estimation procedure for σ leads to intractable equations. In addition, estimating σ by the Weiss' estimate $\hat{\sigma}_n^{(3)}$ is much more simpler than

using the L-estimator $\hat{\sigma}_n^{(1)}$ of (3.23). Now:

$$\hat{\sigma}_n^{(3)} = -\sum_{j=1}^{n-1} \left[\left(1 - \frac{j}{n+1}\right) \log\left(1 - \frac{j}{n+1}\right) \right] (X_{j+1:n} - X_{j:n})$$

and by Theorem (5.1), we have

$$M_n^{(3)}(\lambda) \xrightarrow{D} G^{(3)}(\lambda) \quad , \tag{6.34}$$

where

$$G^{(3)}(\lambda) = \int_0^1 \frac{B^2(y)}{(1-y)\log(1-y)} \left(\log\log\frac{1}{1-y}\right)^{\lambda-1} dy$$

$$+ (T^{3)})^2 J^{(\lambda+1)} - 2\, T^{(3)} R^{(\lambda)} \quad ,$$

where

$$R^{(\lambda)} = \int_0^1 B(y) \left(\log\log\frac{1}{1-y}\right)^{\lambda} dy \quad ,$$

$$J^{(\alpha)} = \int_{-\infty}^{\infty} x^{\alpha} e^{2x} e^{-2e^{x}} dx \quad ,$$

$$T^{(3)} = \int_0^1 \frac{1 + \log(1-y)}{(1-y)\log(1-y)} B(y)\, dy \quad .$$

As far as the Weibull distribution is concerned, i.e., when X has the distribution function $G(x) = 1 - \exp\{-(\frac{x}{c})^{\sigma}\}$. Defining the random variable Z as $Z = \log X$ and denoting c by e^{μ}, we get that the random variable Z has a density function $f \in \mathrm{EXTVAL}(\mu,\sigma)$. Hence the above procedure can be used to test for the Weibull distribution.

CHAPTER III

ON THE ESTIMATED QUANTILE PROCESS: FURTHER GOODNESS-OF-FIT CONSIDERATIONS

1. Introduction

In chapter II, we considered the asymptotic distribution of quadratic nuisance parameter free goodness-of-fit tests for the location and scale parameters family of distribution function $\mathcal{F}$ of (1.1) of chapter II. The methodology used in chapter II can be employed to deduce more results concerning the same family $\mathcal{F}$ of (1.1) of chapter II. In the present chapter we shall consider the problem of weak convergence of the estimated standardized empirical quantile process. In addition, some linear nuisance parameter free goodness-of-fit procedures will be discussed.

For the sake of further reference we repeat in the following few lines some of the basic definitions. We start with the family $\mathcal{F}$ of location and scale parameters distribution functions, i.e.,

$$\mathcal{F} = \{F(x;\mu,\sigma) : F(x;\mu,\sigma) = F_0((x-\mu)/\sigma),\ -\infty < u < +\infty\ ,\ \sigma > 0\}. \quad (1.1)$$

Further let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution function $F \in \mathcal{F}$, with F_0 specified. The empirical quantile function Q_n, based on $X_1, X_2, \ldots, X_n$, is defined as

$$Q_n(y) = \begin{cases} X_{k:n} & \text{if } \frac{k-1}{n+1} < y \leq \frac{k}{n+1} \quad , \; k = 1,2,\ldots,n, \\ X_{n:n} & \text{if } \frac{n}{n+1} < y \leq 1 \quad , \end{cases} \tag{1.2}$$

where $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ is the order statistics of the random sample $X_1, X_2, \ldots, X_n$, and if the continuous function F_0 of (1.1) has a density function f_0, the standardized empirical quantile process of the family F of (1.1), is defined by

$$\begin{aligned} \rho_n(y;\mu,\sigma) &= n^{1/2} f_0(F_0^{-1}(y))((Q_n(y) - \mu)/\sigma - F_0^{-1}(y)) \\ &= n^{1/2} f_0(F_0^{-1}(y))(Q_n^0(y) - F_0^{-1}(y)) \\ &= \rho_n^0(y) \qquad , \; 0 < y < 1 \quad , \end{aligned} \tag{1.3}$$

where $Q_n^0(y) = (Q_n(y) - \mu)/\sigma$ and $Q_n(y)$ is as in (1.2).

Let $\hat{\theta}_n = (\hat{\mu}_n \quad \hat{\sigma}_n)$ be a sequence of estimators for $\theta = (\mu \; \sigma)$ based on the sample $X_1, X_2, \ldots, X_n$, and define the estimated stantardized empirical quantile process $\hat{\rho}_n^0(y)$ as

$$\begin{aligned} \hat{\rho}_n^0(y) &= \rho_n(y;\hat{\mu}_n,\hat{\sigma}_n) \\ &= n^{1/2} f_0(F_0^{-1}(y)) \; \{(Q_n(y) - \hat{\mu}_n)/\hat{\sigma}_n - F_0^{-1}(y)\} \quad , \\ &\qquad 0 < y \leq 1 \quad . \end{aligned} \tag{1.4}$$

In section 2, we will consider the problem of weak convergence of the estimated standardized empirical quantile process $\hat{\rho}_n^0(.)$ of (1.4), when estimating $\theta = (\mu\ \sigma)$ by different sequences of estimators. More precisely, we will consider the following sequences of estimators: $\hat{\theta}_n^{(1)} = (\hat{\mu}_n^{(1)}\ \ \hat{\sigma}_n^{(1)})$, the sequence of maximum likelihood (ML), or maximum likelihood-like (ML-like) estimators; $\hat{\theta}_n^{(2)} = (\hat{\mu}_n^{(2)}\ \ \hat{\sigma}_n^{(2)})$, the sequence of L-estimators (i.e., linear combinations of order statistics); and finally, if μ is known, we will consider Weiss' estimate $\hat{\sigma}_n^{(3)}$ of σ.

Concerning the sequence $\hat{\theta}_n^{(1)}$ of ML or ML-like estimators of θ based on $X_1, X_2, \ldots, X_n$, we assume that (cf. (2.1), (2.2) and Remark (2.1) of chapter II)

(i) $n^{1/2}(\hat{\theta}_n^{(1)} - \theta_0) = n^{-1/2}\sum_1^n \ell(X_j,\theta_0) + \varepsilon_n$,

where $\ell(.,\theta)$ is a measurable 2-dimensional vector valued function, θ_0 is the true value of θ, and $\varepsilon_n \xrightarrow{P} 0$.

(ii) $E\,\ell(X_j,\theta_0) = 0$, $j = 1,2,\ldots,n$. (1.5)

(iii) $E\{\ell(X_j,\theta_0)^t\,\ell(X_j,\theta_0)\}$ is a finite non-negative definite matrix, where A^t is the transpose of the matrix A.

(iv) $\ell_1(x,\theta_0)$, $\ell_2(x,\theta_0)$ are of bounded variation on

each finite interval, where

$$\ell(.,\theta_0) = (\ell_1(.,\theta_0) \qquad \ell_2(.,\theta_0))^t$$

(v) $$\ell(X_j,\theta_0) = \sigma\left(\ell_1\left(\frac{X_j - \mu}{\sigma}\right) \qquad \ell_2\left(\frac{X_j - \mu}{\sigma}\right)\right)^t .$$

(vi) $$\lim_{h\uparrow 1} ((1 - h) \log \log(1 - h)^{-1})^{1/2} \, \|\ell(F^{-1}(h,\theta_0),\theta_0)\| = 0,$$

and $$\lim_{h\downarrow 0}(h \log \log h)^{1/2} \, \|\ell(F^{-1}(h,\theta_0),\theta_0)\| = 0 ,$$

where $\|\cdot\|$ on R^2 is defined by

$$\|(y_1,y_2)\| = \max(|y_1|,|y_2|) .$$

Concerning the sequence $\hat{\theta}_n^{(2)}$ of L-estimators (cf. (3.20) to (3.27) of chapter II) we recall

$$\hat{\theta}_n^{(2)} = (\hat{\mu}_n^{(2)} \qquad \hat{\sigma}_n^{(2)}) \quad , \quad \text{where}$$

$$\left.\begin{aligned} \hat{\mu}_n^{(2)} &= \frac{1}{n} \sum_{j=1}^{n} J_1\left(\frac{j}{n+1}\right) X_{j:n} , \\ \hat{\sigma}_n^{(2)} &= \frac{1}{n} \sum_{j=1}^{n} J_2\left(\frac{j}{n+1}\right) X_{j:n} , \end{aligned}\right\} \qquad (1.6)$$

where , J_1, J_2 are suitable weight functions.

Finally, if μ is assumed to be known, then Weiss' estimate $\hat{\sigma}_n^{(3)}$ of σ is (cf. §5 of chapter II)

$$\hat{\sigma}_n^{(3)} = \frac{1}{n} \sum_{j=1}^{n} f_0(F_0^{-1}(\frac{j}{n+1})) (X_{j+1:n} - X_{j:n}) , \tag{1.7}$$

where $f_0 = F_0'$.

Let F_n be the empirical distribution function based on $X_1, X_2, \ldots, X_n$. The empirical process based on $X_1, X_2, \ldots, X_n$ is defined as

$$\alpha_n(t) = n^{1/2}(F_n(t) - F(t;\mu,\sigma))$$

$$= n^{1/2}(F_n(t) - F_0(\frac{t-\mu}{\sigma})) \tag{1.8}$$

We recall from chapter II

<u>Lemma 3A</u>

Assume that all the conditions of Theorem 1A are satisfied. Then, there exists a sequence of Brownian bridges $\{B_n\}$ for each n , such that

$$\sup_t |\alpha_n(t) - B_n(F_0(\frac{t-\mu}{\sigma}))| \ \underline{\text{a.s.}} \ O(n^{-1/2}\log n) \tag{1.9}$$

and

$$\sup_y |\rho_n^0(y) - B_n(y)| \ \underline{\text{a.s.}} \ O((n^{-1}\log \log n)^{1/4} (\log n)^{1/2}) , \quad (1.10)$$

where $\alpha_n(.)$ is as in (1.8) and $\rho_n^0(.)$ is as in (1.3).

As to the estimated standardized empirical quantile process $\hat{\rho}_n^0(.)$ of the family F of (1.1) we proved:

Theorem 3B

Assume that all the conditions of Theorem 1A are satisfied. Then

(i) Estimating θ by $\hat{\theta}_n^{(1)}$, the sequence of ML or ML-like estimators which satisfy condition (1.5)(i)-(vi), we have

$$\hat{\rho}_n^0(y) = \frac{1}{1 + o_p(1)} \left\{ \rho_n^0(y) - T_n^{(11)} f_0(F_0^{-1}(y)) - T_n^{(12)} f_0(F_0^{-1}(y)) \right.$$

$$\left. F_0^{-1}(y) + o_p(1) \ (f_0(F_0^{-1}(y)) + f_0(F_0^{-1}(y))F_0^{-1}(y)) \right\} , \quad (1.11)$$

(ii) Estimating θ by $\hat{\theta}_n^{(2)}$, the sequence of L-estimators of (1.6) with a suitable choice of J_1, J_2 (cf. Lemma 3.1 of chapter II), and assuming that the standard conditions for the validity of the Cramér-Rao bounds are satisfied, we have

$$\hat{\rho}_n^0(y) = \frac{1}{1 + o_p(1)} \Big\{ \rho_n^0(y) - T_n^{(21)} f_0(F_0^{-1}(y)$$

$$- T_n^{(22)} f_0(F_0^{-1}(y))\ F_0^{-1}(y)$$

$$+ o_p(1)(f_0(F_0^{-1}(y))$$

$$+ f_0(F_0^{-1}(y))\ F_0^{-1}(y)) \Big\} \quad , \qquad (1.2)$$

(iii) Estimating σ by $\hat{\sigma}_n^{(3)}$ (μ is assumed to be known), Weiss' estimate of σ, and assuming that f_0' is continuous on (a,b), we have

$$\hat{\rho}_n^0(y) = \frac{1}{1 + o_p(1)} \Big\{ \rho_n^0(y) - T_n^{(3)} f_0(F_0^{-1}(y))\ F_0^{-1}(y)$$

$$+ o_p(1)\ f_0(F_0^{-1}(y))\ F_0^{-1}(y) \Big\} \quad , \qquad (1.13)$$

where

$$T_n^{(1i)} = \int \ell_i(x)\ d_x\ B_n(F_0(x)) \quad , \quad i = 1,2 \quad , \qquad (1.14)$$

$\ell_i(.)$, $i = 1,2$ are as in (1.5) ,

$$T_n^{(2i)} = \int J_i(x)\, B_n(x)\, d\, F_0^{-1}(x) \quad , \quad i = 1,2 \quad , \tag{1.15}$$

$J_i(.)$, $i = 1,2$ are as in (1.6) and

$$T_n^{(3)} = \int \frac{f_0'(F_0^{-1}(x))}{f_0^2(F_0^{-1}(x))}\, B_n(x)\, dx \quad . \tag{1.16}$$

2. On the Weak Convergence of the Estimated Quantile Process

First we prove a Theorem 1A - type statement for $\hat{\rho}_n^0$ of (1.4).

Theorem 2.1

Let $X_1, X_2, \ldots, X_n$ be a random sample with a continuous distribution function $F \in \mathcal{F}$ of (1.1). Assume that F_0 of (1.1) satisfies the conditions of Theorem 1A as well as

$$\sup_{a<x<b} f_0(x) \leq \text{Constant} \qquad \text{and}$$

$$\sup_{a<x<b} |x| f_0(x) \leq \text{Constant} \; . \tag{2.1}$$

Then

(i) When estimating θ by $\hat{\theta}_n^{(1)}$, the sequence of ML or ML-like estimators of Theorem 3B (i), we have

$$\sup_{0<y<1} \left|\hat{\rho}_n^0(y) - \frac{1}{1 + o_p(1)}\, \hat{B}_n^{(1)}(y)\right| = o_p(1) \; , \tag{2.2}$$

(ii) When estimating θ by $\hat{\theta}_n^{(2)}$, the sequence of L-estimators of Theorem 3B (ii), we have

$$\sup_{0<y<1} \left| \hat{\rho}_n^0(y) - \frac{1}{1 + o_p(1)} \hat{B}_n^{(2)}(y) \right| = o_p(1) \quad , \tag{2.3}$$

(iii) When using Weiss' estimate $\hat{\theta}_n^{(3)}$ of Theorem 3B (iii), we have

$$\sup_{0<y<1} \left| \hat{\rho}_n^0(y) - \frac{1}{1 + 0_p(1)} \hat{B}_n^{(3)}(y) \right| = o_p(1) \quad , \tag{2.4}$$

where $\{\hat{B}_n^{(1)}(y) \; ; \; 0 \leq y \leq 1\}$ is defined in terms of the Brownian bridges $\{B_n\}$ $(n = 1, 2, \ldots)$ of Lemma 3A as follows:

$$\hat{B}_n^{(1)}(y) = B_n(y) - f_0(F_0^{-1}(y))T_n^{(11)} - f_0(F_0^{-1}(y))F_0^{-1}(y)T_n^{(12)} \quad , \tag{2.6}$$

and $\{\hat{B}_n^{(2)}(y)\}$ and $\{\hat{B}_n^{(3)}(y)\}$ are defined in terms of the Brownian bridges $\{B_n\}$ of Theorem 1A as follows

$$\hat{B}_n^{(2)}(y) = B_n(y) - f_0(F_0^{-1}(y))T_n^{(21)} - f_0(F_0^{-1}(y))\, F_0^{-1}(y)T_n^{(22)}, \tag{2.7}$$

$$\hat{B}_n^{(3)}(y) = B_n(y) - f_0(F_0^{-1}(y))\, F_0^{-1}(y)T_n^{(3)} \quad , \tag{2.8}$$

with $T_n^{(ij)}$, $i = 1,2$, $j = 1, 2$ are as in (1.16) and (1.17) and $T_n^{(3)}$ is as in (1.18).

Proof

We will only prove (2.2), since the proofs of (2.3), (2.4) are similar to that of (2.2).

By (1.14) and (2.1)

$$\sup_{0<y<1}\left|\hat{\rho}_n^0(y) - \frac{1}{1+o_p(1)}\left\{\rho_n^0(y) - f_0(F_0^{-1}(y))T_n^{(11)} - f_0(F_0^{-1}(y))\,F_0^{-1}(y)T_n^{(12)}\right\}\right| = o_p(1)\ . \tag{2.9}$$

By Lemma 3A

$$\sup_{0<y<1}\left|\rho_n^0(y) - B_n(y)\right| \ \underline{\underline{a.s.}}\ O((n^{-1}\log\log n)^{1/4}(\log n)^{1/2}) \tag{2.10}$$

Combining (2.9) and (2.10) gives (2.2).

Remark 1

Theorem (2.1) tells us that the estimated standardized quantile process $\hat{\rho}_n^0(.)$ of the family F of (1.1) is near to a Gaussian process which is nuisance parameter free (i.e., free of the unknown parameters μ and σ of the family F), but not distribution free (i.e., it depends on the assumed to be known form of F_0 of F of (1.1).

Corollary 2.1

(i) When estimating θ by $\hat{\theta}_n^{(1)}$, the sequence of ML or ML-like estimators of Theorem 3B (i), we have

$$\sup_{0<y<1}|\hat{\rho}_n^0(y)| \xrightarrow{D} \sup_{0<y<1}|B(y) - f_0(F_0^{-1}(y))T^{(11)}$$

$$- f_0(F_0^{-1}(y))\, F_0^{-1}(y)T^{(12)}|, \quad (2.11)$$

(ii) When estimating θ by $\hat{\theta}_n^{(2)}$, the sequence of L-estimators of Theorem 3B (ii), we have

$$\sup_{0<y<1}|\hat{\rho}_n^0(y)| \xrightarrow{D} \sup_{0<y<1}|B(y) - f_0(F_0^{-1}(y))T^{(21)}$$

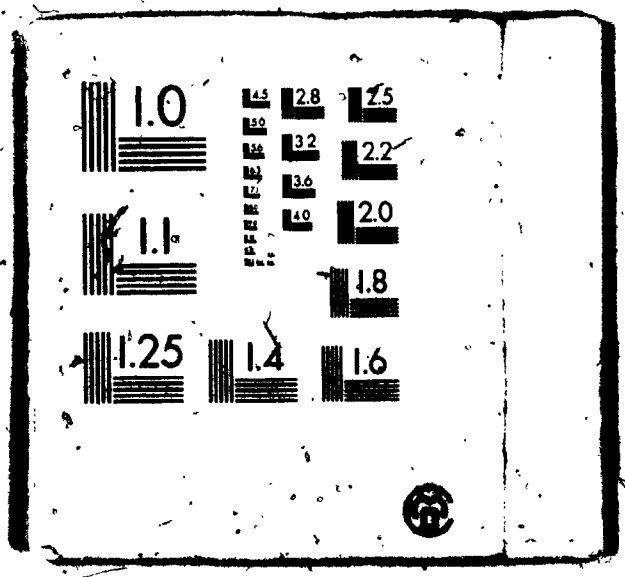
$$- f_0(F_0^{-1}(y))\, F_0^{-1}(y)T^{(22)}|, \quad (2.12)$$

(iii) When estimating σ by $\hat{\sigma}_n^{(3)}$, Weiss' estimate of Theorem 3B (iii), we have

$$\sup_{0<y<1}|\hat{\rho}_n^0(y)| \xrightarrow{D} \sup_{0<y<1}|B(y) - f_0(F_0^{-1}(y))F_0^{-1}(y)T^{(3)}|, \quad (2.13)$$

where

2 OF / DE 2



$$T^{(1i)} = \int \ell_i(x)\, d_x\, B(F_0(x)) \quad , \; i = 1,2 \quad , \tag{2.14}$$

$$T^{(2i)} = \int J_i(x)\, B\,(x)\, d\, F_0^{-1}(x) \quad , \; i = 1,2 \quad , \tag{2.15}$$

$$T^{(3)} = \int \frac{f_0'(F_0^{-1}(y))}{f_0^2(F_0^{-1}(y))}\; B(y)\, dy \quad \ldots \tag{2.16}$$

Considering the composite goodness-of-fit hypothesis

$$H_0 : F \in \mathcal{F} \text{ with } F_0 \text{ specified,} \tag{2.17}$$

the test statistics of the left hand sides of (2.11), (2.12) and (2.13) and the distribution of the right hand side r.v.'s, can be used, at least in principle, to test for H_0 of (2.17).

Now as a possible alternative route to the problem of testing for H_0 of (2.17), we consider next the following linear form of $\hat{\rho}_n^0$ of (1.4):

$$L_n(\lambda) = \sum_{k=1}^{n} \hat{\rho}_n^0\left(\frac{k}{n+1}\right) \cdot \left(F_0^{-1}\left(\frac{k}{n+1}\right)\right)^{\lambda-1} / n f_0\left(F_0^{-1}\left(\frac{k}{n+1}\right)\right) ,$$

$$\lambda = 1; 2 \ldots \tag{2.18}$$

The considerations we have in mind when considering $L_n(\lambda)$ are similar to those which led to introduce $M_n(\lambda)$ of (1.16) in Csörgő, Révész (1979, 1981)

3. On the Asymptotic Distribution of $L_n(\lambda)$ of (2.18)

Theorem 3.1

Let $X_1, X_2, \ldots, X_n$ be a random sample with a continuous distribution function $F \in \mathcal{F}$ of (1.1). Assume that F_0 of $\mathcal{F}$ satisfies all the conditions of Theorem 1A as well as $|E_0 X^\lambda| < +\infty$ and

$$\lim_{y\to 0} y^{1/r} |F_0^{-1}(y)| = \lim_{y\to 1} (1-y)^{1/r} |F_0^{-1}(y)| = 0 \tag{3.1}$$

with $r > 2(\lambda+1)$ $(\lambda > 1, 2, \ldots)$. Assume further that $L_n^{(i)}(\lambda)$ is that version of $L_n(\lambda)$, when θ is estimated via $\hat{\theta}_n^{(i)}$, $i = 1, 2, 3$. Then

(i) Under the same conditions as in Theorem 3B (i), we have

$$\left| L_n^{(1)}(\lambda) - \left\{ I_n^{(\lambda-1)} - T_n^{(11)} \int_0^1 (F_0^{-1}(y))^{\lambda-1}\, dy - T_n^{(12)} \int_0^1 (F_0^{-1}(y))^{\lambda}\, dy \right\} \right| = o_p(1) \tag{3.2}$$

and

$$L_n^{(1)}(\lambda) \xrightarrow{D} I^{(\lambda-1)} - T^{(11)} \int_0^1 (F_0^{-1}(y))^{\lambda-1}\, dy - T^{(12)}$$

$$\int_0^1 (F_0^{-1}(y))^{\lambda}\, dy = L^{(1)}(\lambda)$$

$$, \quad \lambda = 1,2,\ldots \quad , \tag{3.3}$$

(ii) Under the same conditions of Theorem 3B (ii), we have

$$\left| L_n^{(2)}(\lambda) - \left\{ I_n^{(\lambda-1)} - T_n^{(21)} \int_a^b x^{\lambda-1} f_0(x)dx - T_n^{(22)} \int_a^b x^{\lambda} f_0(x)dx \right\} \right| = o_p(1) \quad , \tag{3.4}$$

and

$$L_n^{(2)}(\lambda) \xrightarrow{D} I^{(\lambda-1)} - T^{(21)} E_0 X^{\lambda-1} - T^{(22)} E_0 X^{\lambda}$$

$$= L^{(2)}(\lambda) \quad , \qquad \lambda = 1,2\ldots \quad , \tag{3.5}$$

(iii) Under the same conditions of Theorem 3B (iii), we have

$$\left| L_n^{(3)}(\lambda) - \left\{ I_n^{(\lambda-1)} - T_n^{(3)} E_0 X^{\lambda} \right\} \right| = o_p(1) \quad , \tag{3.6}$$

and

$$L_n^{(3)}(\lambda) \xrightarrow{D} I^{(\lambda-1)} - T^{(3)}E_0X^{\lambda} = L^{(3)}(\lambda) \qquad \lambda = 1,2,\ldots, \tag{3.7}$$

where

$$I_n^{(\lambda-1)} = \int_0^1 B_n(y)\lambda^{-1}\, d(F_0^{-1}(y))^{\lambda}, \quad \lambda = 1,2,\ldots, \tag{3.8}$$

$$I^{(\lambda-1)} = \int_0^1 B(y)\lambda^{-1}\, d(F_0^{-1}(y))^{\lambda}, \quad \lambda = 1,2,\ldots, \tag{3.9}$$

$$E_0X^{\lambda} = \int_a^b x^{\lambda} f_0(x)dx \tag{3.10}$$

$T_n^{(1i)}$, $i = 1, 2$, $T_n^{(2i)}$, $i = 1, 2$, and $T_n^{(3)}$ are as in (1.16), (1.17) and (1.18) respectively, and $T^{(1i)}$, $i = 1, 2$, $T^{(2i)}$, $i = 1, 2$ and $T^{(3)}$ are as in (2.14), (2.15) and (2.16) respectively.

<u>Proof</u>

We will only prove (3.4), since the proofs of (3.2) and (3.6) are similar. By (1.12), and (2.18), we have

$$L_n^{(2)}(\lambda) = \left\{ \sum_{k=1}^{n} \frac{\rho_n^0(\frac{k}{n+1})}{n\, f_0(F_0^{-1}(\frac{k}{n+1}))} (F_0^{-1}(\frac{k}{n+1}))^{\lambda-1} \right.$$

$$- T_n^{(21)} \sum_{k=1}^{n} \frac{(F_0^{-1}(\frac{k}{n+1}))^{\lambda-1}}{n}$$

$$- T_n^{(22)} \sum_{k=1}^{n} \frac{(F_0^{-1}(\frac{k}{n+1}))^{\lambda}}{n}$$

$$+ o_p(1) \left\{ \sum_{k=1}^{n} \frac{(F_0^{-1}(\frac{k}{n+1}))^{\lambda-1}}{n} \right.$$

$$\left. + \sum_{k=1}^{n} \frac{(F_0^{-1}(\frac{k}{n+1}))^{\lambda}}{n} \right\} / (1 + o_p(1))$$

$$\lambda = 1,2,\ldots \qquad (3.11)$$

Now, by condition (3.1) with $r > 2(\lambda+1)$ $(\lambda = 1, 2, \ldots)$ we have

$$\left| \sum_{k=1}^{n} \frac{(F_0^{-1}(\frac{k}{n+1}))^{\lambda-1}}{n} - \int_0^1 (F_0^{-1}(y))^{\lambda-1}\, dy \right| =$$

$$= \left| \sum_{k=1}^{n} \frac{(F_0^{-1}(\frac{k}{n+1}))^{\lambda-1}}{n} - E_0 X^{\lambda-1} \right| = o_p(n^{-1/2}) ,$$

$$\lambda = 1,2,\ldots \qquad (3.12)$$

By a proof similar to that of Theorem 2A (Csörgő, Révész (1981)) we can easily prove that

$$\left| \sum_{k=1}^{n} \frac{\rho_n^0(\frac{k}{n+1})}{n\, f_0(F_0^{-1}(\frac{k}{n+1}))} \left(F_0^{-1}(\frac{k}{n+1})\right)^{\lambda-1} - I_n^{(\lambda-1)} \right| = o_p(1) \quad (3.13)$$

Combining (3.1), (3.12) and (3.13) we get (3.4).

As to a test statistic for H_0 of (2.17) in terms of $L_n^{(i)}(\lambda)$, we should reject H_0 if $|L_n^{(i)}(\lambda)|$ is too large, and since by (3.3), (3.5) and (3.7) we have

$$|L_n^{(i)}(\lambda)| \xrightarrow{D} |L^{(i)}(\lambda)| \quad , \quad i = 1,2,3 \; ; \; \lambda = 1,2,\ldots , \quad (3.14)$$

a tabulation of $|L^{(i)}(\lambda)|$ is theoretically feasible for some specific F_0 (say, for the standard normal distribution, i.e., when $F_0 = \Phi$). For each given F_0 of F of (1.1) the asymptotic distribution of $|L_n^{(i)}(\lambda)|$ will depend on F_0; i.e., the proposed test procedure for H_0 of (2.17) is nuisance parameter free but non-distribution free.

A rationalization of $L_n^{(i)}(\lambda)$ for H_0 of (2.17) is along the lines of that of $M_n^{(i)}(\lambda)$ of chapter II. The difficulties one is facing when attempting to tabulate the distribution of $L_n^{(i)}(\lambda)$ are somewhat less formidable than those in case of $M_n^{(i)}(\lambda)$. This is one of the reasons for proving Theorem (3.1). As to the relative merits of any one of the statistics $L_n^{(i)}(\lambda)$ and $M_n^{(i)}(\lambda)$ when testing for H_0 of (2.17) via them, at this stage we can only say that specific case-studies will only tell which one of these procedures is to be preferred

in any given practical situation. We observe, however, that if λ is an odd integer, and if H_0 of (2.17) is not true then $|L_n^{(i)}(\lambda)|$ a.s. ∞ as $n \to \infty$, $i = 1, 2, 3$, i.e., just like the $M_n^{(i)}(\lambda)$ test in this case, the proposed test is then also consistent against any alternative hypothesis.

The techniques of the previous sections can be used in goodness-of-fit tests for a wide class of distributions. Indeed, as examples, we may go through the examples of section 6 of chapter II, i.e., on testing for the exponential, the logistic, extreme value and the Weibull distributions. However, in order to avoid repeating ourself, in the next section we will only discuss goodness-of-fit tests for the exponential distribution.

4. On Testing for Exponentiality

Consider the family of exponential density functions.

$$\mathrm{EXP}(A,B) = \{f(x;A,B) : f(x;A,B) = f_0(\frac{x - A}{B}) =$$

$$= B^{-1}\exp(-(x - A)/B), \quad x \geq A \in R^1, \ B > 0\} \quad . \qquad (4.1)$$

Let $X_1, X_2, \ldots, X_n$ be a random sample with a density function f, we wish to test the following composite null hypotheses

$$H_0^{(1)} : f \in \mathrm{EXP}(0,\sigma) \quad , \quad \sigma > 0 \quad , \qquad (4.2)$$

and

$$H_0^{(2)} : f \in \mathrm{EXP}(\mu,\sigma) \quad , \quad \mu \in R^1 \quad , \sigma > 0 \qquad (4.3)$$

First we consider $H_0^{(1)}$ of (4.2). Under $H_0^{(1)}$, $X_1, X_2, \ldots, X_n$ is a random sample with density function $f \in EXP(0,\sigma)$, $\sigma > 0$. As in section 6 of chapter II, we consider three alternative routes.

Case 1

Using the ML estimate of σ i.e., $\hat{\sigma}_n^{(1)} = \bar{X}_n$ consequently

$$L_n(\lambda) = \sum_{k=1}^{n} n^{-1/2} \left\{ \frac{X_{k:n}}{\bar{X}_n} - \log \frac{1}{1 - \frac{k}{n+1}} \right\} \left(\log \frac{1}{1 - \frac{k}{n+1}} \right)^{\lambda-1} \tag{4.4}$$

A similar argument as that used in section 6 of chapter II, and Theorem (3.1) will give

$$L_n^{(1)}(\lambda) \xrightarrow{D} L^{(1)}(\lambda) \quad , \qquad \lambda = 1,2,\ldots \quad , \tag{4.5}$$

where

$$L^{(1)}(1) = 0 \quad , \qquad \text{and}$$

$$L^{(1)}(\lambda) = \int_0^1 \frac{B(y)}{1-y} \left(\log \frac{1}{1-y}\right)^{\lambda-1} dy - \Gamma(\lambda+1) \int_0^1 \frac{B(y)}{1-y}\, dy \quad , \quad \lambda = 2,3,\ldots \tag{4.6}$$

For example if $\lambda = 2$, then we have

$$L_n^{(1)}(2) \to \int_0^1 \frac{B(y)}{1-y} \left(\log \frac{1}{1-y}\right) dy - 2\int_0^1 \frac{B(y)}{1-y}\, dy \quad , \tag{4.7}$$

and we should reject $H_0^{(1)}$ of (4.2) if $|L_n^{(1)}(2)|$ is too large. A similar statement holds for $L_n^{(1)}(\lambda)$ of (4.4) in general.

An alternative procedure is via Theorem (2.1), as follows:

$$\hat{\rho}_n^0\left(\frac{k}{n+1}\right) = n^{1/2}\left(1 - \frac{k}{n+1}\right)\left\{\frac{X_{k:n}}{\bar{X}_n} - \log\frac{1}{1 - \frac{k}{n+1}}\right\}, \tag{4.8}$$

and by Theorem (2.1), we have

$$\hat{\rho}_n^0(.) \xrightarrow{D} B(.) - (1 - .)\log\frac{1}{1 - .}\int_0^1 \frac{B(y)}{1 - y}\,dy, \tag{4.9}$$

where $\{B(y)\ ;\ 0 \le y \le 1\}$ is a Brownian bridge.

Whence for $\hat{\rho}_n^0$ of (4.8) we have, for example, that

$$\max_{1\le k\le n} \left|\hat{\rho}_n^0\left(\frac{k}{n+1}\right)\right| \xrightarrow{D}$$

$$\sup_{0<y<1}\left|B(y) - (1 - y)\log\frac{1}{1 - y}\int_0^1 \frac{B(x)}{1 - x}\,dx\right|, \tag{4.10}$$

and a test for $H_0^{(1)}$ of (4.2) can be based on large values of

$$\max_{1\le k\le n} \left|\hat{\rho}_n^0\left(\frac{k}{n+1}\right)\right| =$$

$$= n^{1/2}\max_{1\le k\le n}\left|\frac{X_{k:n}}{\bar{X}_n} - \log\frac{1}{1 - \frac{k}{n+1}}\right|\left(1 - \frac{k}{n+1}\right). \tag{4.11}$$

Case 2

Using the L-estimator of σ of $EXP(0,\sigma)$, which, in this case, is $\hat{\sigma}_n^{(2)} = \bar{X}_n$ = M.L.E. Hence, in this case, $L_n^{(1)}(\lambda) = L_n^{(2)}(\lambda)$. An argument similar to that of case 2 of section 6 of chapter II, will give (4.5) through (4.11).

Case 3

Using Weiss' estimate of σ, namely

$$\hat{\sigma}_n^{(3)} = \frac{1}{n+1}(n\bar{X}_n + X_{n:n} - (n+1)X_{1:n}). \tag{4.12}$$

Now

$$L_n^{(3)}(\lambda) = \sum_{k=1}^{n} n^{1/2} \left\{ \frac{X_{k:n}}{\hat{\sigma}_n^{(3)}} - \log \frac{1}{1 - \frac{k}{n+1}} \right\} \left\{ \log \frac{1}{1 - \frac{k}{n+1}} \right\}^{\lambda-1}$$

$$\xrightarrow{D} L^{(3)}(\lambda) \overset{D}{=} L^{(1)}(\lambda) \quad , \quad \lambda = 1,2,\ldots , \tag{4.13}$$

where $L^{(1)}(\lambda)$ is as in (4.6).

In testing for $H_0^{(2)}$ of (4.3), one can make the following transformation: $X_i^* = X_i - X_{1:n}$, $i = 1, 2, \ldots, n$ and delete from it the term which equals to zero, say X_n^* (with probability one there is one and only one such term). Let $X_{1:n}^*, X_{2:n}^*, \ldots, X_{n-1:n}^*$ be the order statistic of $X_1^*, \ldots, X_{n-1}^*$. Given $H_0^{(2)}$ of (4.3), $X_1^*, \ldots, X_{n-1}^*$

are independent EXP(0,σ)rv's, and the above procedure thus modified, can now be used to test for $H_0^{(2)}$ of (4.3).

CHAPTER IV

ON EXPONENTIAL AND UNIFORM SPACINGS

1. Introduction and Preliminaries

Let $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ be the order statistics of a random sample $X_1, X_2, \ldots, X_n$ from some distribution function F. The spacings T_j's are then defined by

$$T_j = X_{j+1:n} - X_{j:n} \quad , \qquad j = 1,2,\ldots,n-1 \quad . \tag{1.1}$$

If the random variable X has a lower bound, A say, and/or an upper bound, B say, then one may define $X_{0:n} = A$ and/or $X_{n+1:n} = B$ and

$$T_0 = X_{1:n} - X_{0:n} \qquad \text{and/or} \qquad T_n = X_{n+1:n} - X_{n:n} \quad . \tag{1.2}$$

When X has uniform distribution, we say that the $\{T_j\}_{j=0}^{n}$ are uniform spacings, whereas, when X has exponential distribution, we say that $\{T_j\}_{j=0}^{n-1}$ are exponential spacings.

In the case of uniform spacings there are a number of works written on the asymptotic distribution of the random variables $G_n = g(T_0, T_1, \ldots, T_n)$, where g is a measurable function. In particular, Moran (1947) considered the asymptotic distribution of $\sum_{j=0}^{n} T_j^2$, Kimball (1950) proved that

$$n^{1/2}\left(n^{\alpha-1} \sum_{m} T_j^{\alpha} - \mu \frac{n^{\alpha}\Gamma(n+1)\,\Gamma(1+\alpha)}{\Gamma(n+\alpha+1)}\right) \quad \text{was}$$

asymptotically normal with mean zero and variance

$$u(\Gamma(2\alpha + 1) - (1 + u\alpha^2)\,\Gamma^2(1 + \alpha)) \quad ,$$

where $\sum_m$ denotes the sum of an arbitrary m of the T_j's , $u = \lim_{n\to\infty} \frac{m}{n} >$ and $\alpha > 0$, $\alpha \neq 1$ if $u = 1$. Also Darling (1953) proved that

$$n^{1/2}(n^{\alpha-1} \sum_{j=0}^{n} T_j^{\alpha} - \Gamma(1 + \alpha))$$

was asymptocially normal with mean zero and variance

$$\Gamma(2\alpha + 1) - (1 + \alpha^2)\,\Gamma^2(1 + \alpha) \quad .$$

Up to date, as far as we know, nothing has been done concerning the asymptotic theory of the stochastic processes $G_n^{\alpha}(u) = g(T_0, T_1, \ldots$ $\ldots, T_{[nu]})$, where $[t]$ is the greatest integer in t and $0 \leq u \leq 1$. Also Pyke (1965 and 1972) pointed out that a study of rates of convergence for limiting distributions of spacings was also desirable. Some answers to these open questions are given in the present chapter. Our main result in this direction is Theorem (2.1), which states that uniformly in $u \in (0,1)$ and for large n

$$n^{1/2}(n^{\alpha-1} \sum_m T_j^{\alpha} - u\,\Gamma(1 + \alpha)) \overset{D}{=} K_n(u;\alpha) + O_{a.s.}(n^{-1/2}\log n), \quad (1.3)$$

where $\sum_m$ denotes the sum of an arbitrary m units of the summands and $u = \lim_{n\to\infty} \frac{m}{n} > 0$, where $K_n(u;\alpha)$, as defined by (2.11) and (2.14) in

the sequel, converges in distribution to

$$K(u;\alpha) \overset{D}{=} \lambda_1(\alpha)\, W(u) + \lambda_2(\alpha)\, B(u) ,$$

where $W(u)$ is a standard Wiener process, $B(u) = W(u) - u\, W(1)$ is a Brownian bridge,

$$\lambda_1(\alpha) = \sqrt{\Gamma(2\alpha + 1) - (1 + \alpha^2)\, \Gamma^2(1 + \alpha)} \quad \text{and}$$

$$\lambda_2(\alpha) = -\lambda_1(\alpha) \pm \sqrt{\lambda_1^2(\alpha) + \alpha^2 \Gamma^2(1 + \alpha)} .$$

An important feature of Theorem (2.1) is that it provides a methodology for finding the asymptotic distribution and rates of convergence of processes of the form

$$p_n(u;\alpha) = \begin{cases} 0 & \text{if } u < \frac{2}{n} \\ n^{1/2} \left(\dfrac{n^{\alpha-1} \sum_{j=1}^{[nu]-1} T_j^{\alpha}}{\left(\sum_{j=1}^{n-1} T_j \right)^{\alpha}} - u\, \Gamma(1 + \alpha) \right) & \text{if } \frac{2}{n} \le u \le 1 , \end{cases} \tag{1.4}$$

$$\sum_{m} h(T_j) \tag{1.5}$$

and

$$\sum_{j=1}^{[nu]-1} h\left(\frac{T_j}{\sum_{1}^{n-1} T_j} \right) \tag{1.6}$$

for a wide class of functions h, as given by Theorem (2.2), Theorem (3.1) and Theorem (3.2) respectively.

In the case of exponential spacings, Theorem (2.3) gives the asymptotic distribution together with a rate of convergence of the stochastic processes $\{J_n(u;\alpha)\ ;\ u \in (0,1]\ ,\ n = 1, 2, \ldots\ ;\ \alpha > 0\}$ defined as

$$J_n(u;\alpha) = \begin{cases} 0 & \text{if } u \le \frac{2}{n} \\ n^{1/2}\left[\dfrac{q_n(u;\alpha)}{(q_n(1;1))^{\alpha}} - u\,\Gamma(1+\alpha)\right] & \text{if } \frac{2}{n} \le u \le 1 \end{cases} \tag{1.7}$$

where

$$q_n(u;\alpha) = n^{\alpha-1} \sum_{j=1}^{[nu]-1} \{f_0(F_0^{-1}(\tfrac{j}{n}))\ T_j\}^{\alpha} ,$$

with $F_0(x) = 1 - e^{-x}$, $f_0(x) = F_0'(x)$ and $\alpha > 0$. In Theorem (3.3), we carry out the same programme for the process

$$\sum_{j=1}^{[nu]-1} h\left[\frac{f_0(F_0^{-1}(\frac{j}{n}))T_j}{q_n(1;1)}\right] \tag{1.8}$$

for a wide class of functions h.

We note that $p_n(u;\alpha)$ of (1.4) and $J_n(u;\alpha)$ of (1.7) are, respectively, the uniform and exponential versions of $g_n(u;\alpha)$ of

(1.6) of chapter II. Consequently, as it is suggested in chapter II, a number of nuisance parameter free goodness-of-fit procedures can be based on $p_n(u;\alpha)$ of (1.4) and $J_n(u;\alpha)$ of (1.7) for the uniform respectively exponential distributions. These procedures can be applied to tests for the random distribution of events in time. This is the story of section 4.

In section 5, we concentrate on the problem of rates of convergence.

The following Theorem will be used in the sequel

<u>Theorem 4A</u> (Komlós, Major, and Tusnády (1975, 1976)).

Let $V_1, V_2, \ldots$ be i.i.d.rv's, the moment generating function of which is finite in some neighbourhood of the origin and assume that $EV_1 = 0$ and $EV_1^2 = 1$. If the probability space is rich enough, then there exists a standard Wiener process W such that for all real z,

$$P\{\max_{1\le k\le n} \left|\sum_{j=1}^{k} V_j - W(k)\right| > (A_4 \log n) + z\} < A_5\, e^{-A_6 z} \tag{1.9}$$

where A_4, A_5 and A_6 are absolute constants. This inequality, in turn, implies

$$\max_{1\le k\le n} \left|\sum_{j=1}^{k} V_j - W(k)\right| \ \underline{\text{a.s.}}\ O(\log n) \tag{1.10}$$

In the sequel it will be always assumed that the underlying probability space is rich enough in the sense of Theorem 4A.

2. Limit Theorems for Sums of Powers of Uniform and Exponential Spacings

Let $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ be the order statistics of a random sample $X_1, X_2, \ldots, X_n$ from the uniform distribution on $[0,1]$. Take $X_{0:n} = 0$ and $X_{n+1:n} = 1$ and define the uniform spacings $\{T_j\}_{j=0}^{n}$ as follows:

$$T_j = X_{j+1:n} - X_{j:n} \quad , \qquad j = 0,1,2,\ldots,n \quad . \tag{2.1}$$

Assume that $\alpha > 0$ is given and consider the stochastic process

$$\{n^{\alpha-1} \sum_{j=0}^{[nu]} T_j^{\alpha} \quad ; \quad u \in (0,1], n = 1,2,\ldots\} \quad . \tag{2.2}$$

In order to study the asymptotic theory of the latter, we let $Z_1, Z_2, \ldots, Z_{n+1}$ be i.i.d.rv's with distribution function $F_0(x) = 1 - e^{-x}$. Then it is well known that, for each n,

$$\{S_k/S_{n+1} ; \ 1 \le k \le n\} \overset{D}{=} \{X_{k:n} \ ; \ 1 \le k \le n\} \quad , \tag{2.3}$$

where $S_0 = 0$ and $S_k = \sum_{j=1}^{k} Z_j$, $k = 1, 2, \ldots, n+1$. Consequently we have for each n

$$\{Z_{j+1}/S_{n+1} \ ; \ 0 \le j \le n\} \overset{D}{=} \{T_j \ ; \ 0 \le j \le n\} \quad . \tag{2.4}$$

Thus, in particular, we have for each n

$$n^{\alpha-1} \sum_{j=0}^{[nu]} T_j^{\alpha} \overset{D}{=} n^{\alpha-1} \sum_{j=0}^{[nu]} (Z_{j+1}/S_{n+1})^{\alpha} = \left(\frac{1}{n} \sum_{j=1}^{[nu]+1} Z_j^{\alpha}\right) / \left(\frac{1}{n} \sum_{j=1}^{n+1} Z_j\right)^{\alpha} ,$$

$$0 < u \le 1 \quad . \tag{2.5}$$

This suggests that one may first study the asymptotic theory of the process on the right hand side of (2.5) in order to derive some asymptotic results for the process on the left hand side of (2.5).

It is well known that the expected value, EZ_1^α, and the variance, $Var(Z_1^\alpha)$, of Z_1^α are given by

$$\begin{cases} EZ_1^\alpha = \Gamma(1+\alpha) \\ Var(Z_1^\alpha) = \Gamma(2\alpha+1) - \Gamma^2(1+\alpha) = V^2(\alpha) \quad \text{, say} \end{cases} \tag{2.6}$$

Hence

$$\frac{1}{n}\sum_{j=1}^{[nu]+1} Z_j^\alpha - \Gamma(1+\alpha)\frac{[nu]+1}{n}$$

$$= \frac{V(\alpha)}{n}\sum_{j=1}^{[nu]+1}\left(\frac{Z_j^\alpha - \Gamma(1+\alpha)}{V(\alpha)}\right) , \tag{2.7}$$

where the right hand side is $\frac{V(\alpha)}{n}$ times a partial sum of i.i.d.rv's with mean zero and variance one. Moreover the moment generating function of the summands exists. Therefore, by applying Theorem 4A (Komlós, Major, Tusnády (1975, 1976)), there exists a standard Wiener process $\{W_{(\alpha)}(t) \; ; \; 0 \le t\}$ such that

$$\sup_{0<u<1}\left|\frac{1}{n}\sum_{j=1}^{[nu]+1} Z_j^\alpha - u\,\Gamma(1+\alpha) - \frac{W_{(\alpha)}(V^2(\alpha)nu)}{n}\right|$$

$$\underline{a.s.} \quad O(n^{-1}\log n) \quad . \tag{2.8}$$

In particular

$$\sup_{0<u<1}\left|\frac{1}{n}\sum_{j=1}^{[nu]+1} Z_j - u - \frac{W_{(1)}(nu)}{n}\right| \underline{\underline{a.s.}}\ O(n^{-1}\log n)\ , \quad (2.9)$$

and consequently,

$$\frac{1}{n}\sum_{j=1}^{n+1} Z_j = 1 + \frac{W_{(1)}(n)}{n} + O_{a.s.}(n^{-1}\log n) \quad . \quad (2.10)$$

It should be noticed that, for any u, $0 \le u \le 1$, we have

$$E\,\frac{1}{\sqrt{n}}\,W_{(\alpha)}\,(V^2(\alpha)\,nu)\,\frac{1}{\sqrt{n}}\,W_{(1)}(n)$$

$$= \lim_{n\to\infty} n\,E\,\{\frac{1}{n}\sum_{j=1}^{[nu]+1}(Z_j^{\alpha} - \Gamma(1+\alpha)\}\,\{\frac{1}{n}\sum_{j=1}^{n+1}(Z_j - 1)$$

$$= \lim_{n\to\infty}\frac{1}{n}\,\{([nu]+1)(EZ_j^{\alpha+1} - \Gamma(1+\alpha)EZ_j)\}$$

$$= \lim_{n\to\infty}\frac{[nu]+1}{n}\,\{\Gamma(2+\alpha) - \Gamma(1+\alpha)\}$$

$$= \alpha\,u\,\Gamma(1+\alpha) \quad . \quad (2.11)$$

Now, by (2.10), we have

$$(\frac{1}{n}\sum_{j=1}^{n+1} Z_j)^{\alpha}\ \underline{\underline{a.s.}}\ \left\{(1 + O(n^{-1}\log n)) + \frac{W_{(1)}(n)}{n}\right\}^{\alpha}$$

$$\underline{\underline{\text{a.s.}}} \quad (1 + O(n^{-1}\log n))^{\alpha} \left\{ 1 + \frac{W_{(1)}(n)}{n(1 + O(n^{-1}\log n))} \right\}^{\alpha}$$

$$\underline{\underline{\text{a.s.}}} \quad (1 + O(n^{-1}\log n)) \left\{ 1 + \frac{\alpha W_{(1)}(n)}{n(1 + o(1))} + O(n^{-1}) \right\}$$

$$\underline{\underline{\text{a.s.}}} \quad 1 + \alpha \frac{W_{(1)}(n)}{n} + O(n^{-1}\log n) \tag{2.12}$$

By (2.8), (2.10), (2.11) and (2.12) we have

$$n^{1/2} \left[\frac{\left(\frac{1}{n} \sum_{j=1}^{[nu]+1} Z_j^{\alpha}\right)}{\left(\frac{1}{n} \sum_{j=1}^{n+1} Z_j\right)^{\alpha}} - u\,\Gamma(1 + \alpha) \right]$$

$$= \frac{n^{1/2}}{\left(\frac{1}{n} \sum_{j=1}^{n+1} Z_j\right)^{\alpha}} \left\{ \left(\frac{1}{n} \sum_{j=1}^{[nu]+1} Z_j^{\alpha}\right) - u\,\Gamma(1 + \alpha) \left(\frac{1}{n} \sum_{j=1}^{n+1} Z_j\right)^{\alpha} \right\}$$

$$\underline{\underline{\text{a.s.}}} \quad \frac{n^{1/2}}{1 + o(1)} \left\{ (u\,\Gamma(1 + \alpha) + \frac{W_{(\alpha)}(V^2(\alpha)\, nu)}{n} + O(n^{-1}\log n)) \right.$$

$$\left. - u\,\Gamma(1 + \alpha)(1 + \alpha \frac{W_{(1)}(n)}{n} + O(n^{-1}\log n)) \right\}$$

$$\underline{\underline{a.s.}}\ \frac{1}{1+o(1)}\left\{ n^{-1/2} W_{(\alpha)}(V^2(\alpha)\ nu) - u\alpha\ \Gamma(1+\alpha) n^{-1/2} W_{(1)}(n) + O(n^{-1/2}\log n)\right\}$$

$$\underline{\underline{a.s.}}\ \frac{1}{1+o(1)}\left\{K_n(u;\alpha) + O(n^{-1/2}\log n)\right\}, \tag{2.13}$$

where

$$K_n(u;\alpha) = \frac{1}{n^{1/2}}\left(W_{(\alpha)}(V^2(\alpha)\ nu) - u\alpha\ \Gamma(1+\alpha)W_{(1)}(n)\right). \tag{2.14}$$

This proves our first result:

Lemma 2.1

Let $Z_1, Z_2, \ldots, Z_{n+1}$ be i.i.d.rv's with exponential distribution function $F_0(x) = 1 - e^{-x}$. Then, uniformly in $u \in (0,1)$, and for n large

$$n^{1/2}\left\{\left(n^{\alpha-1}\sum_{j=1}^{[nu]+1} Z_j^{\alpha}\right)/\left(\frac{1}{n}\sum_{j=1}^{n+1} Z_j\right)^{\alpha} - u\ \Gamma(1+\alpha)\right\}$$

$$\underline{\underline{a.s.}}\ K_n(u;\alpha) + O(n^{-1/2}\log n) \tag{2.15}$$

where $K_n(u;\alpha)$ is as in (2.14).

Now we go back to our original aim, namely, to study the asymptotic theory of the processes of (2.2). By (2.15) and Lemma (2.1), we get immediately

Theorem 2.1

Let $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ be the order statistics of a random sample $X_1, X_2, \ldots, X_n$ from the uniform distribution on $[0,1]$. Let the spacings T_j, $j = 0, 1, \ldots, n$ be as defined in (2.1). Then, uniformly in $u \in (0,1)$, and for n large

$$n^{1/2}\left(n^{\alpha-1}\sum_{j=0}^{[nu]} T_j^{\alpha} - u\,\Gamma(1+\alpha)\right) \overset{D}{=} K_n(u;\alpha) + O_{a.s.}(n^{-1/2}\log n) \quad . \tag{2.16}$$

Remark 2.1

It is clear that

$$K_n(u;\alpha) \xrightarrow{D} K(u;\alpha) = W_{(\alpha)}(V^2(\alpha)\,u) - u\alpha\,\Gamma(1+\alpha)W_{(1)}(1) \quad , \tag{2.17}$$

where

$$EK(u;\alpha) = 0 \quad ,$$

and (cf. (2.11))

$$E\,K(u_1;\alpha)\,K(u_2;\alpha) = \{\Gamma(2\alpha+1) - (1+\alpha^2)\,\Gamma^2(1+\alpha)\}\,u_1 \wedge u_2 + \alpha^2\Gamma^2(1+\alpha)\,\{u_1 \wedge u_2 - u_1 u_2\} \quad . \tag{2.19}$$

Also, it is clear that

$$K(u;\alpha) \overset{D}{=} \lambda_1(\alpha)\, W(u) + \lambda_2(\alpha)\, B(u) \quad , \tag{2.20}$$

where $W(u)$ is a standard Wiener process, and, in terms of the latter, $B(u) = W(u) - u\, W(1)$, a Brownian bridge,

$$\lambda_1(\alpha) = \{\Gamma(2\alpha + 1) - (1 + \alpha^2)\, \Gamma^2(1 + \alpha)\}^{1/2}$$

and

$$\lambda_2(\alpha) = \lambda_1(\alpha) \pm \{\lambda_1^2(\alpha) + \alpha^2 \Gamma^2(1 + \alpha)\}^{1/2} \quad .$$

Corollary 2.1

Using the notation of Theorem (2.1), we have, uniformly in $u \in (0,1)$,

$$n^{1/2}(n^{\alpha-1} \sum_{j=0}^{[nu]} T_j^{\alpha} - u\, \Gamma(1 + \alpha)) \overset{D}{\to} \lambda_1(\alpha)\, W(u) + \lambda_2(\alpha)\, B(u) \quad , \tag{2.21}$$

where $W(u)$, $B(u)$, $\lambda_1(\alpha)$ and $\lambda_2(\alpha)$ are as in (2.20).

Corollary 2.2

For $\alpha > 0$, $\alpha \neq 1$, we have for each n large enough

$$n^{1/2}(n^{\alpha-1} \sum_{j=0}^{n} T_j^{\alpha} - \Gamma(1 + \alpha)) \overset{D}{=} K_n(1;\alpha) + O_{a.s.}(n^{-1/2} \log n) \quad , \tag{2.22}$$

where $K_n(1;\alpha) \overset{D}{\to} K(1;\alpha) \overset{D}{=} \lambda_1(\alpha)\, W(1)$, a normal rv with mean zero and variance $\lambda_1^2(\alpha) = (2\alpha + 1) - (1 + \alpha^2)\, \Gamma^2(1 + \alpha)$. This

agrees with Kimball's (1950) and Darling's (1953) results.

Now, Theorem (2.1) can be proved in a more general form, namely

Theorem (2.1)*

Using the notation of Theorem (2.1), we have, uniformly in $u \in (0,1)$ and for n large

$$n^{1/2}(n^{\alpha-1} \sum_m T_j^\alpha - u\,\Gamma(1+\alpha)) \overset{D}{=} K_n(u;\alpha) + O_{a.s.}(n^{-1/2}\log n) \quad , \qquad (2.23)$$

where $\sum_m T_j^\alpha$ denotes the sum of m units arbitrary chosen from the $(n+1)$ units $T_0^\alpha, T_1^\alpha, \ldots, T_n^\alpha$, and $u = \lim_{n\to\infty} \frac{m}{n} > 0$.

Proof

One need only to notice that for each n

$$n^{\alpha-1} \sum_m T_j^\alpha \overset{D}{=} (\frac{1}{n} \sum_m Z_j^\alpha)/(\frac{1}{n} \sum_{j=1}^{n+1} Z_j) \quad , \qquad (2.24)$$

where the Z_j's are i.i.d.rv's with distribution function $F_0(x) = 1 - e^{-x}$ and $\sum_m Z_j^\alpha$ is the sum of m i.i.d.rv's each with mean $\Gamma(1+\alpha)$ and variance $V^2(\alpha)$ and each has a moment generating function.

A useful consequence of Theorems (2.1) and (2.1)* is the following:

Theorem 2.2

Let $X_{1:n}, \ldots, X_{n:n}$ be the order statistics of a random sample $X_1, X_2, \ldots, X_n$ from the uniform distribution on $[0,1]$.

Let the spacings T_j , $j = 1, 2, \ldots, n-1$, be as defined in (2.1). Then, uniformly in $u \in (0,1)$, and for n large

$$p_n(u;\alpha) \overset{D}{=} K_n(u;\alpha) + O_{a.s.}(n^{-1/2}\log n) \quad , \tag{2.25}$$

where $p_n(u;\alpha)$ is as in (1.4). Consequently (cf. (2.17) and (2.20)), uniformly in $u \in (0,1)$

$$p_n(u;\alpha) \overset{D}{\to} \lambda_1(\alpha)\, W(u) + \lambda_2(\alpha)\, B(u) \quad \text{as} \quad n\to\infty \quad . \tag{2.26}$$

Remark 2.2

Given the random sample $Y_1, Y_2, \ldots, Y_n$ from the uniform distribution on $[A,B]$ for some $A < B \in R^1$ we let $X_i = \dfrac{Y_i - A}{B - A}$, $i = 1, 2, \ldots, n$. Then $X_1, X_2, \ldots, X_n$ is a random sample from the uniform distribution on $[0,1]$. Now, if $Y_{1:n}, Y_{2:n}, \ldots, Y_{n:n}$ are the order statistics of the random sample $Y_1, \ldots, Y_n$ and $X_{1:n}, \ldots, X_{n:n}$ are the order statistics of the sample $X_1, X_2, \ldots, X_n$, then

$$n^{1/2}\left[\frac{n^{\alpha-1}\sum_{1}^{[nu]-1}(Y_{j+1:n} - Y_{j:n})^{\alpha}}{\left(\sum_{1}^{n-1}(Y_{j+1:n} - Y_{j:n})\right)^{\alpha}} - u\,\Gamma(1+\alpha)\right]$$

$$\overset{D}{=} n^{1/2}\left[\frac{n^{\alpha-1}\sum_{1}^{[nu]-1}(X_{j+1:n} - X_{j:n})^{\alpha}}{\left(\sum_{1}^{n-1}(X_{j+1:n} - X_{j:n})\right)^{\alpha}} - u\,\Gamma(1+\alpha)\right] \quad . \tag{2.27}$$

Hence $p_n(u;\alpha)$ of (1.4) is nuisance parameter free and the statements of Theorem (2.2) is true provided that the random sample $X_1, X_2, \ldots, X_n$ in the Theorem is coming from a uniform distribution on some finite interval.

Concerning exponential spacings, let $T_1, T_2, \ldots, T_{n-1}$ be exponential spacings from $F_0(x) = 1 - e^{-x}$ and consider the stochastic processes $\{J_n(u;\alpha)\ ;\ u \in (0,1],\ n = 1, 2, \ldots,\ \alpha > 0\}$ as defined in (1.7). It is well known that $(n - j)T_j\ ;\ j = 1, 2, \ldots, n-1$ are i.i.d.rv's with distribution function $F_0(x) = 1 - e^{-x}$. So that by knowing that $f_0(F_0^{-1}(\frac{j}{n}))T_j = \frac{(n-j)}{n} T_j$ and by Theorem (1.1), we now have

Theorem 2.3

In the case of exponential spacings $T_1, T_2, \ldots, T_{n-1}$ from $F_0(x) = 1 - e^{-x}$, we have

$$\sup_{0<u<1} |J_n(u;\alpha) - K_n(u;\alpha)| \ \underline{\text{a.s.}}\ O(n^{-1/2}\log n) \quad , \tag{2.28}$$

where $J_n(u;\alpha)$ is as in (1.7).

Corollary 2.3

Let $\{T_j\}_{j=1}^{n-1}$ be exponential spacings from the exponential distribution $F(x) = F_0(\frac{x-\mu}{\sigma})$, $F_0(x) = 1 - e^{-x}$, μ and σ are nuisance parameters. Then for each real number $\alpha > 0$ we have (2.28).

Remark 2.3

The special cases $\alpha = 1$, $0 < u < 1$ of (2.25) and of (2.28) are also special cases of a general result of Csörgő, Révész (1980) and their result agrees with ours. As to (2.16), the case $\alpha = 2$, $u \neq 1$ was examined before by Moran (1947), the case $u = 1$, $\alpha > 0$ was examined by Darling (1953) and Koziol (1980) and the case $0 < u < 1$, $\alpha > 0$ was partially studied by Kimball (1950). Our result (2.16) agrees with all the above mentioned results. In addition (2.16) gives a solution to the problem under consideration in a closed form together with the indicated rate of convergence.

Remark 2.4

The special case $\alpha = 1$, $0 < u < 1$ of (2.28) is also a special case of a general result of Csörgő, Révész (1980) and their result agrees with ours. The special case $\alpha = 2$, $0 < u \leq 1$ of (2.28) was first suggested by Weiss (1957) for a general class of distributions. He did not, however, prove any asymptotic distribution results.

3. The Asymptotic Distributions of $\sum h(T_j)$

Let $T_0, T_1, \ldots, T_n$ be uniform spacings on $[0,1]$ (as defined in (2.1)). Suppose that

$$h(x) = \sum_{r=0}^{s} a_r x^r \quad , \quad s \geq 2 \quad . \tag{3.1}$$

For the moment suppose that $a_1 \neq 0$ and a_ℓ is the first non-zero coefficient after a_1. We consider the asymptotic distribution of the process

$$\sum_{j=0}^{[nu]} h(T_j) = \sum_{j=0}^{[nu]} \left(\sum_{r=0}^{s} a_r T_j^r \right)$$

$$= \sum_{r=0}^{s} a_r \left(\sum_{j=0}^{[nu]} T_j^r \right)$$

$$= a_0([nu] + 1) + \sum_{r=1}^{s} a_r \left(\sum_{j=0}^{[nu]} T_j^r \right) . \tag{3.2}$$

In particular, at $u = 1$, we have

$$\sum_{j=0}^{n} h(T_j) = (n + 1)a_0 + a_1 + \sum_{r=\ell\geq 2}^{s} a_r \left(\sum_{j=0}^{s} T_j^r \right) . \tag{3.3}$$

By Theorem (2.1), we have

$$n^{\ell-1} \sum_{j=0}^{n} h(T_j) \overset{D}{=} n^{\ell-1}(n + 1)a_0 + n^{\ell-1}a_1 + n^{\ell-1} \sum_{r=\ell}^{s} a_r \left\{ \frac{r!}{n^{r-1}} \right.$$

$$\left. + \frac{K_n(1;r)}{n^{r-\frac{1}{2}}} + 0_{a.s.}(n^{-r}\log n) \right\}$$

$$= n^{\ell-1}(n+1)a_0 + n^{\ell-1}a_1 + \ell!a_\ell + \sum_{r=\ell+1}^{s} \frac{a_r r!}{n^{r-\ell}}$$

$$+ a_\ell n^{-1/2} K_n(1;\ell) + \sum_{r=\ell+1}^{s} \frac{a_r}{n^{r-\ell+\frac{1}{2}}} K_n(1;r)$$

$$+ O_{a.s.}(n^{-\ell}\log n) \quad .$$

Hence

$$n^{1/2}(n^{\ell-1} \sum_{j=0}^{n} h(T_j) - (n^{\ell-1}(n+1)a_0 + n^{\ell-1}a_1 + \ell!a_\ell))$$

$$\stackrel{D}{=} a_\ell K_n(1;\ell) + O_{a.s.}(n^{-1/2}\log n) \tag{3.4}$$

Consequently (cf. (2.17) and (2.20)), the rv on the left hand side of (3.4) converges in distribution to

$$a_\ell \lambda_1(\ell) W(1) \quad ,$$

a normal rv with mean zero and variance $a_\ell^2 \lambda_1^2(\ell)$, where $\lambda_1(\cdot)$ is as defined in (2.20).

Now, suppose that $a_1 = 0$ and a_ℓ , $\ell \geq 2$ is the first non-zero coefficient after a_1 . By applying Theorem (2.1) to the terms on the right hand side of (3.2), we get

$$n^{\ell-1}\sum_{j=0}^{[nu]} h(T_j) \overset{D}{=} n^{\ell-1}\left[nua_0 + \sum_{r=\ell}^{s} a_r \left\{\frac{r!}{n^{r-1}}u + \frac{K_n(u;r)}{n^{r-\frac{1}{2}}} + O_{a.s.}(n^{-r}\log n)\right\}\right]$$

$$= n^{\ell}a_0u + \ell!a_\ell u + a_\ell n^{-1/2} K_n(u;\ell)$$

$$+ \sum_{r=\ell+1}^{s} \frac{a_r}{n^{r-\ell+\frac{1}{2}}} K_n(u;r) + O_{a.s.}(n^{-\ell}\log n) .$$

Therefore, uniformly in $u \in (0,1)$ and for n large

$$n^{1/2}(n^{\ell-1}\sum_{j=0}^{[nu]} h(T_j) - n^{\ell}a_0u - \ell!a_\ell u)$$

$$\overset{D}{=} a_\ell K_n(u;\ell) + O_{a.s.}(n^{-1/2}\log n) \quad . \tag{3.5}$$

Consequently, (cf. (2.17) and (2.20)), uniformly in $u \in (0,1)$

$$n^{1/2}(n^{\ell-1}\sum_{j=0}^{[nu]} h(T_j) - n^{\ell}a_0u - \ell!a_\ell u) \overset{D}{\to} a_\ell\lambda_1(\alpha)\, W(u)$$

$$+ a_\ell\lambda_2(\ell)\, B(u) \quad , \tag{3.6}$$

where $W(u)$, $B(u)$, $\lambda_1(.)$ and $\lambda_2(.)$ are as in (2.20).

Finally, we consider the case $a_1 \neq 0$ and $0 < u < 1$. By Theorem (2.1) and (3.2), we have

$$\sum_{j=0}^{[nu]} h(T_j) \overset{D}{=} nua_0 + ua_1 + a_1 \frac{K_n(u;1)}{n^{1/2}} + \sum_{r=2}^{s} a_r \frac{r!u}{n^{r-1}}$$

$$+ \sum_{r=2}^{s} \frac{a_r K_n(u;r)}{n^{r-\frac{1}{2}}} + O_{a.s.}(n^{-1}\log n) \quad . \tag{3.7}$$

And hence, uniformly in $u \in (0,1)$ and for large n, we have

$$n^{1/2}\left(\sum_{j=0}^{[nu]} h(T_j) - nua_0 - ua_1\right) \overset{D}{\to} a_1(\lambda_1(1)\, W(u) + \lambda_2(1)\, B(u)$$

$$= a_1\, B(u) \quad . \tag{3.8}$$

This completes the proof of the following

<u>Lemma 3.1</u>

Let $\{T_j\}_{j=0}^{n}$ be uniform spacings taken from the uniform distribution on $[0,1]$. For $h(.)$ of (3.1), we have

(i) If $a_1 \neq 0$ and a_ℓ, $\ell \geq 2$, is the first non-zero coefficient after a_1, then (3.4) holds true.

(ii) If $a_1 \neq 0$, then uniformly in $u \in (0,1)$ and for large n we have (3.7). Also, (3.8) holds true.

(iii) If $a_1 = 0$ and a_ℓ, $\ell \geq 2$ is the first non-zero coefficient after a_1, then uniformly in $u \in (0,1)$ and for n large we have (3.5). Also, uniformly in $u \in (o,1)$ (3.6) holds true.

Corollary 3.1

Let $h(.)$ be any continuous real valued function which has a Taylor expansion

$$h(x) = \sum_{r=0}^{\infty} a_r x^r \quad , \quad |x| \leq 1 \quad . \tag{3.9}$$

Then the statement of Lemma (3.1) holds true for $h(.)$ of (3.9) in place of $h(.)$ of (3.1).

Naturally, for functions $h(.)$ other than (3.1) and (3.9), one should look for a minimum set of regularity conditions to be imposed on $h(.)$ in order to formulate a Lemma (3.1) - like result in terms of $\sum_{j=0}^{[nu]} h(T_j)$.

Let us start with the following set of assumptions:

$$\left.\begin{array}{ll} \text{(i)} & h(.) \text{ is twice differentiable with bounded second derivative on } [0,1] \\ \text{(ii)} & h'(0) \neq 0 \end{array}\right\} \tag{3.10}$$

In this case we have

$$h(x) = h(0) + h'(0)x + h''(c)\frac{x^2}{2} \quad , \quad 0 < c < x \tag{3.11}$$

and concentrating only on the last term with $0 < c_j < T_j$, $j = 0, 1, \ldots, n$, we consider

$$\left| \sum_{j=0}^{[nu]} \frac{h''(c_j)}{2} T_j^2 \right| \le \frac{1}{2} \max_{0 \le j \le n} |h''(c_j)| \sum_{j=0}^{[nu]} T_j^2$$

$$\le \frac{1}{2} \max_{0 \le x \le 1} |h''(x)| \sum_{j=0}^{[nu]} T_j^2$$

$$\stackrel{D}{=} \frac{1}{2} \max_{0 \le x \le 1} |h''(x)| \left\{ \frac{2u}{n} + \frac{K_n(u;2)}{n^{3/2}} + O_{a.s.}(n^{-2} \log n) \right\}$$

$$\underline{\underline{a.s.}} \; O(n^{-1} \log n) \quad . \tag{3.12}$$

Therefore, uniformly in $u \in (0,1)$ and for large n

$$n^{1/2} \left(\sum_{j=0}^{[nu]} h(T_j) - nuh(0) - uh'(0) \right)$$

$$\stackrel{D}{=} h'(0) \, K_n(u;1) + O_p(n^{-1/2} \log n) \quad . \tag{3.13}$$

Consequently (cf. (2.17) and (2.20)), uniformly in $u \in (0,1)$

$$n^{1/2} \left(\sum_{j=0}^{[nu]} h(T_j) - nuh(0) - uh'(0) \right) \stackrel{D}{\to} h'(0) \, B(u) \quad . \tag{3.14}$$

Now, if condition (ii) of (3.10) does not hold, then instead of (3.10), we assume the following set of assumptions:

$$\left.\begin{array}{l}\text{(i) } h \text{ is } (\ell+1) \text{ differentiable, where } \ell \text{ is the} \\ \quad\text{first positive integer such that } h^{(\ell)} \neq 0 . \\ \text{(ii) } h^{(\ell+1)}(.) \text{ is bounded on } [0,1] .\end{array}\right\} \qquad (3.10)^*$$

In this case, we have

$$h(x) = h(0) + \frac{h^{(\ell)}(0)}{\ell!} x^{\ell} + \frac{h^{(\ell+1)}(c)}{(\ell+1)!} x^{\ell+1}, \quad 0 < c < x . \qquad (3.12)^*$$

Then again, we can easily prove that uniformly in $u \in (0,1)$ and n large

$$n^{1/2}(n^{\ell-1} \sum_{j=0}^{[nu]} h(T_j) - n^{\ell} uh(0) - uh^{(\ell)}(0))$$

$$\stackrel{D}{=} \frac{h^{(\ell)}(0)}{\ell!} R_n(u;\ell) + O_p(n^{-1/2} \log n) \qquad (3.15)$$

Consequently (cf. (2.17) and (2.20)), uniformly in $u \in (0,1)$

$$n^{1/2}(n^{\ell-1} \sum_{j=0}^{[nu]} h(T_j) - n^{\ell} uh(0) - uh^{(\ell)}(0))$$

$$\stackrel{D}{\to} \frac{h^{(\ell)}(0)}{\ell!} (\lambda_1(\ell) W(u) + \lambda_2(\ell) B(u)) . \qquad (3.16)$$

Therefore, we have

Theorem 3.1

Let $T_0, T_1, \ldots, T_n$ be the uniform spacings on $[0,1]$ and let $h(\cdot)$ be a real valued function which satisfies condition (3.10)*. Then uniformly in $u \in (0,1)$, (3.15) for n large and (3.16) hold true.

Remark 3.1

One disadvantage of $\sum_{j=0}^{[nu]} h(T_j)$ is that $h(0) \neq 0$, but if we can choose the function $h(.)$ such that $h(0) = 0$, then we would have

Corollary 3.2

If in addition to all the conditions of Theorem (3.1), we also have $h(0) = 0$, then uniformly in $u \in (0,1)$ and n large

$$n^{1/2}(n^{\ell-1} \sum_{j=0}^{[nu]} h(T_j) - uh^{(\ell)}(0))$$

$$\stackrel{D}{=} \frac{h^{(\ell)}(0)}{\ell!} K_n(u;\ell) + O_p(n^{-1/2}\log n) \quad , \tag{3.17}$$

and (cf. (2.17) and (2.20)), uniformly in $u \in (0,1)$

$$n^{1/2}(n^{\ell-1} \sum_{j=0}^{[nu]} h(T_j) - uh^{(\ell)}(0))$$

$$\stackrel{D}{\to} \frac{h^{(\ell)}(0)}{\ell!} (\lambda_1(\ell)\, W(u) + \lambda_2(\ell)\, B(u)) \quad . \tag{3.18}$$

Remark 3.2

Another disadvantage of $\sum_{j=0}^{[nu]} h(T_j)$, is that it is not nuisance parameter free. To overcome this disadvantage we consider the following Theorem.

Theorem 3.2

Let $T_0, T_1, \ldots, T_n$ be the uniform spacings on $[0,1]$ and let $h(.)$ be a real valued function which satisfies condition (3.10)*. Then uniformly in $u \in (0,1)$ and for n large

$$n^{1/2}\left(n^{\ell-1} \sum_{j=1}^{[nu]-1} h\left(\frac{T_j}{\sum_1^{n-1} T_j}\right) - un^{\ell}h(0) - uh^{(\ell)}(0)\right)$$

$$\stackrel{D}{=} \frac{h^{(\ell)}(0)}{\ell!} K_n(u;\ell) + O_p(n^{-1/2}\log n) \quad , \tag{3.19}$$

and (cf. (2.17) and (2.20)), uniformly in $u \in (0,1)$

$$n^{1/2}\left(n^{\ell-1} \sum_{j=1}^{[nu]-1} h\left(\frac{T_j}{\sum_1^{n-1} T_j}\right) - un^{\ell}h(0) - uh^{(\ell)}(0)\right)$$

$$\stackrel{D}{\to} \frac{h^{(\ell)}(0)}{\ell!} (\lambda_1(\ell)\, W(u) + \lambda_2(\ell)\, B(u)) \tag{3.20}$$

If in addition, we have $h(0) = 0$, then uniformly in $u \in (0,1)$ and for large n

$$n^{1/2}\left(n^{\ell-1}\sum_{j=1}^{[nu]-1} h\left(\frac{T_j}{\sum_{1}^{n-1} T_j}\right) - uh^{(\ell)}(0)\right)$$

$$\stackrel{D}{=} \frac{h^{(\ell)}(0)}{\ell!} K_n(u;\ell) + O_p(n^{-1/2}\log n) \quad . \tag{3.21}$$

and, uniformly in $u \in (0,1)$

$$n^{1/2}\left(n^{\ell-1}\sum_{j=1}^{[nu]-1} h\left(\frac{T_j}{\sum_{1}^{n-1} T_j}\right) - uh^{(\ell)}(0)\right)$$

$$\stackrel{D}{\to} \frac{h^{(\ell)}(0)}{\ell!} (\lambda_1(\ell)\, W(u) + \lambda_2(\ell)\, B(u)) \quad . \tag{3.22}$$

The proof of Theorem (3.2) is similar to that of Theorem (3.1), except that Theorem (2.2) will be used in place of Theorem (2.1), which has been used in proving Theorem (3.1).

Remark 3.3

By the same argument of Remark (2.2), we can easily see that $h\left(\frac{T_j}{\sum_{j=1}^{n-1} T_j}\right)$, $j = 1, 2, \ldots, n-1$ are nuisance parameter free. Hence the statement of Theorem (3.2) is true provided that the random sample $X_1, X_2, \ldots, X_n$ in the Theorem is coming from a uniform distribution

on some finite interval.

Imitating now Theorem (3.2), one can also state and prove

Theorem 3.3

Let $T_1, T_2, \ldots, T_{n-1}$ be exponential spacings from the exponential distribution function $F(x) = F_0(\frac{x-\mu}{\sigma})$, $F_0(x) = 1 - e^{-x}$, $-\infty < \mu < \infty$, $\sigma > 0$ are nuisance parameters. Let $h(.)$ be a real valued function which satisfies condition $(3.10)^*$, then

$$\sup_{0<u<1} \left| n^{1/2} \left\{ n^{\ell-1} \sum_{j=1}^{[nu]-1} h\left(\frac{f_0(F_0^{-1}(\frac{j}{n}))T_j}{q_n(1;1)}\right) - u(n^{\ell} h(0) + h^{(\ell)}(0)) \right\} - \frac{h^{(\ell)}(0)}{\ell!} K_n(u;\ell) \right|$$

$$\underline{a.s.} \quad O(n^{-1/2} \log n) \qquad (3.23)$$

If in addition, we have $h(0) = 0$, then

$$\sup_{0<u<1} \left| n^{1/2} \left\{ n^{\ell-1} \sum_{j=1}^{[nu]-1} h\left(\frac{f_0(F_0^{-1}(\frac{j}{n}) T_j}{q_n(1;1)}\right) - uh^{(\ell)}(0) \right\} - \frac{h^{(\ell)}(0)}{\ell!} K_n(u;\ell) \right|$$

$$\underline{a.s.} \quad O(n^{-1/2} \log n) \qquad , \qquad (3.24)$$

where $q_n(1;1)$ is as in (1.7).

Proof

The proof is similar to that of Theorem (3.1), except that Theorem (2.3) will be used in place of Theorem (2.1), which has been used in proving Theorem (3.1).

4. Nuisance Parameter Free Goodness-Of-Fit Tests for the Uniform and Negative Exponential Families

The main task of this section is to introduce some new test statistics, based on uniform and exponential spacings, together with their asymptotic distributions. In addition, we will examine some known ones. We start with test statistics based on uniform spacings.

We consider the family of uniform density functions

$$U(A,B) = \{f(x;A,B) : f(x;A,B) = (B-A)^{-1} \quad , A,B \in R^1, A < B\} \quad . \tag{4.1}$$

Let $X_1, X_2, \ldots, X_n$ $(n \geq 2)$ be i.i.d.rv with a density function f and, on the basis of this random sample, we wish to test the following composite null hypothesis

$$H_0^{(1)} : f \in U(A,B) \qquad ; \qquad A,B \in R^1 \quad , \quad A < B \quad . \tag{4.2}$$

Now, under $H_0^{(1)}$, we have by Theorem (2.2), Remark (2.1) and Remark (2.2) that, uniformly in $u \in (0,1)$,

$$P_n(u;\alpha) \xrightarrow{D} \lambda_1(\alpha)\, W(u) + \lambda_2(\alpha)\, B(u) \quad , \tag{4.3}$$

where $p_n(u;\alpha)$ is as defined in (1.4) and $\lambda_1(\alpha)\ W(u) + \lambda_2(\alpha)\ B(u)$ is as defined in (2.20). Then in the light of (4.3), we also have

$$\Psi(p_n(u;\alpha)) \xrightarrow{D} \Psi(\lambda_1(\alpha)\ W(u) + \lambda_2(\alpha)\ B(u)) \tag{4.4}$$

as $n \to \infty$, for every continuous functional Ψ on the space of real valued functions on $[0,1]$ endowed with the supremum topology. In particular, we have that

$$t_1(p,n) = p_n(1;\alpha) \xrightarrow{D} \lambda_1(\alpha)\ W(1) \quad , \tag{4.5}$$

$$t_2(p,n) = \sup_{0<u<1} |p_n(u;\alpha)|$$

$$\xrightarrow{D} \sup_{0<u<1} |\lambda_1(\alpha)\ W(u) + \lambda_2(\alpha)\ B(u)| \quad , \tag{4.6}$$

and

$$t_3(p,n) = \int_0^1 p_n^2(u;\alpha)du \quad .$$

$$\xrightarrow{D} \int_0^1 (\lambda_1(\alpha)\ W(u) + \lambda_2(\alpha)\ B(u))^2 du \quad . \tag{4.7}$$

As to testing for $H_0^{(1)}$ of (4.2), we should reject $H_0^{(1)}$ when $|t_1(p,n)|$ (alternatively, $t_2(p,n)$ or $t_3(p,n)$) is too large whatever the value of α might be.

As to testing for exponentiality, consider the family of exponential density functions

$$EXP(A,B) = \{f(x;A,B) : f(x;A,B) = f_0(\frac{x - A}{B}) = B^{-1}(\exp(-(x - A)/B) ,$$

$$x \geq A \in R^1 \quad , \quad B > 0\} \quad . \tag{4.8}$$

Let $X_1, X_2, \ldots, X_n$ $(n > 2)$ be i.i.d.rv with a density function f and, in the light of this random sample, we wish to test the following composite goodness-of-fit null hypothesis

$$H_0^{(2)} : f \in EXP(A,B) \quad , \quad A \in R^1 \quad , \quad B > 0 \tag{4.9}$$

Assuming $H_0^{(2)}$, then by Theorem (2.3) we have, uniformly in $u \in (0,1)$

$$J_n^v(u;\alpha) \xrightarrow{D} \lambda_1(\alpha)\, W(u) + \lambda_2(\alpha)\, B(u) \quad , \tag{4.10}$$

where $J_n(u;\alpha)$ is as in (1.7) and $\lambda_1(\alpha)\, W(u) + \lambda_2(\alpha)\, B(u)$ is as (2.20). Then (4.10) imply

$$\Psi(J_n(u;\alpha)) \xrightarrow{D} \Psi(\lambda_1(\alpha)\, W(u) + \lambda_2(\alpha)\, B(u)) \quad . \tag{4.11}$$

as $n \to \infty$, for every continuous functional Ψ on the space of real valued functions on $[0,1]$ endowed with the supremum topology. In particular, we have that

$$t_1(J,n) = J_n(1;\alpha) \xrightarrow{D} \lambda_1(\alpha)\, W(1) \quad , \tag{4.12}$$

$$t_2(J,n) = \sup_{0<u<1} |J_n(u;\alpha)|$$

$$\xrightarrow{D} \sup_{0<u<1} |\lambda_1(\alpha)\, W(u) + \lambda_2(\alpha)\, B(u)| \quad , \tag{4.13}$$

and

$$t_3(J,n) = \int_0^1 J_n^2(u;\alpha)du$$

$$\xrightarrow{D} \int_0^1 (\lambda_1(\alpha)\ W(u) + \lambda_2(\alpha)\ B(u))^2 du \qquad (4.14)$$

As to testing for $H_0^{(2)}$ of (4.9), we may use any one of $t_i(J,n)$, $i = 1, 2, 3$, and we should reject $H_0^{(2)}$ when $|t_1(J,n)|$ (respectively $t_2(J,n)$, $t_3(J,n)$) is too large, whatever the value of α might be.

5. On the Rates of Convergence

In the present section we consider the problem of rates of convergence. Our discussion will be confined to the case $\alpha = 2$ for two reasons. First, as shown in Sethuraman and Rao (1970), the statistics $n^{\alpha-1}\sum_1^n T_j^\alpha$ has the maximum Pitman efficiency if $\alpha = 2$. Second, the case $\alpha = 2$ is relatively simple to handle. An important step is the following theorem.

Theorem 5.1

Let $Z_1, Z_2, \ldots, Z_n$ be i.i.d.rv with distribution function $F_0(x) = 1 - e^{-x}$. If $K_n(u;\alpha)$ is as in (2.14), then

$$P\left\{\sup_{0<u\le 1}\left|n^{1/2}\left[\frac{\frac{1}{n}\sum_{1}^{[nu]} Z_j^2}{(\frac{1}{n}\sum_{1}^{n} Z_j^2)^2} - 2u\right] - K_n(u;2)\right|\right.$$

$$\left. > \frac{L\log n + x}{n^{1/2}}\right\} \le D\, e^{-\delta x} + \frac{8\quad e^{-x/32}}{(2\pi(L\log n + x))^{1/2}\, n^{L/32}}$$

$$+ \sqrt{2/\pi}\,\frac{1}{\varepsilon\, n^{1/2}}\; e^{-\varepsilon^2 n/2} \qquad , \quad (5.1)$$

where $\varepsilon > 0$ is arbitrary, L, D, δ are positive absolute constants and $x \in R^1$, $x \le 32n - \log n$.

<u>Proof</u>

By Theorem 4A (Komlós, Major, Tusnády (1976), for all $x \in R^1$ we have

$$P\left\{\sup_{0<u\le 1}\left|\frac{1}{n}\sum_{j=1}^{[nu]} Z_j^{\alpha} - \left(u\,\Gamma(1+\alpha) + \frac{W_{(\alpha)}(V^2(\alpha)\,nu)}{n}\right)\right|\right.$$

$$\left. > \frac{A\log n + x}{n}\right\} \le B\, e^{-Cx} \qquad , \quad \alpha = 1,2 \qquad , \quad (5.2)$$

where A, B and C are positive absolute constants.

Now,

$$P = P\left\{\sup_{0<u\leq 1}\left|n^{1/2}\left(\frac{\frac{1}{n}\sum_{1}^{[nu]} Z_j^2}{(\frac{1}{n}\sum_{1}^{n} Z_j)^2} - 2u\right) - K_n(u;\alpha)\right| > \frac{L\log n + x}{n^{1/2}}\right\}$$

$$= P\left\{\sup_{0<u\leq 1}\left|n^{1/2}\left(\frac{\frac{1}{n}\sum_{1}^{[nu]} Z_j^2}{(\frac{1}{n}\sum_{1}^{n} Z_j)^2} - 2u\right) - n^{1/2}\left(\frac{1}{n}\sum_{1}^{[nu]} Z_j^2\right.\right.\right.$$

$$\left.\left. - 2u\left(\frac{1}{n}\sum_{1}^{n} Z_j\right)^2\right) + n^{1/2}\left(\frac{1}{n}\sum_{1}^{[nu]} Z_j^2 - 2u\left(\frac{1}{n}\sum_{1}^{n} Z_j\right)^2\right)\right.$$

$$\left. - K_n(u;2)\right| > \frac{L\log n + x}{n^{1/2}}\Bigg\}$$

$$= P\left\{\sup_{0<u\leq 1}\left|\left(\frac{\frac{1}{n}\sum_{1}^{[nu]} Z_j^2}{(\frac{1}{n}\sum_{1}^{n} Z_j)^2} - 2u\right)\left(1 - \left(\frac{1}{n}\sum_{1}^{n} Z_j\right)^2\right)\right.\right.$$

$$\left.\left. + \left(\frac{1}{n}\sum_{1}^{[nu]} Z_j^2 - 2u\left(\frac{1}{n}\sum_{1}^{n} Z_j\right)^2 - \frac{K_n(u;2)}{n^{1/2}}\right)\right| > \frac{L\log n + x}{n}\right\}$$

$$\leq P\left\{\left|1-\left(\frac{1}{n}\sum_1^n Z_j\right)^2\right| \sup_{0<u\leq 1}\left|\frac{\frac{1}{n}\sum_1^{[nu]} Z_j}{\left(\frac{1}{n}\sum_1^n Z_j\right)^2}-2u\right| > \frac{L\log n + x}{2n}\right\}$$

$$+ P\left\{\sup_{0<u\leq 1}\left|\frac{1}{n}\sum_1^{[nu]} Z_j^2 - 2u\left(\frac{1}{n}\sum_1^n Z_j\right)^2 - \frac{K_n(u;2)}{n^{1/2}}\right| > \frac{L\log n + x}{2n}\right\},$$

or

$$P \leq P_1 + P_2 \tag{5.3}$$

As to P_2 of (5.3), we have

$$P_2 = P\left\{\sup_{0<u\leq 1}\left|\frac{1}{n}\sum_1^{[nu]} Z_j^2 - 2u - 2u\left(\left(\frac{1}{n}\sum_1^n Z_j\right)^2 - 1\right)\right.\right.$$

$$\left.\left. - \frac{W_{(2)}(20\ nu)}{n} + 4u\ \frac{W_{(1)}(n)}{n}\right| > \frac{L\log n + x}{2n}\right\}$$

$$= P\left\{\sup_{0<u\leq 1}\left|\left(\frac{1}{n}\sum_1^{[nu]} Z_j^2 - 2u - \frac{W_{(2)}(20\ nu)}{n}\right)\right.\right.$$

$$\left.\left. - 2u\left(\left(\frac{1}{n}\sum_1^n Z_j\right)^2 - \left(1+\frac{2}{n}W_{(1)}(n)\right)\right)\right| > \frac{L\log n + x}{2n}\right\}$$

$$\leq P_{21} + P_{22}, \tag{5.4}$$

where

$$P_{21} = P\left\{\sup_{0<u\leq 1}\left|\frac{1}{n}\sum_{1}^{[nu]} Z_j^2 - 2u - \frac{1}{n}W_{(2)}(20\ nu)\right| > \frac{L\log n + x}{4n}\right\} \quad (5.5)$$

and

$$P_{22} = P\left\{\left|\left(\frac{1}{n}\sum_{1}^{n} Z_j\right)^2 - \left(1 + \frac{2}{n}W_{(1)}(n)\right)\right| > \frac{L\log n + x}{8n}\right\} . \quad (5.6)$$

In order to estimate P_{22} of (5.6), we argue as follows;

$$P_{22} = P\left\{\left|\left(\frac{1}{n}\sum_{1}^{n} Z_j\right)^2 - \left(1 + \frac{1}{n}W_{(1)}(n)\right)^2 + \frac{1}{n^2}W_{(1)}^2(n)\right|\right.$$

$$\left. > \frac{L\log n + x}{8n}\right\} \leq P\left\{\left|\left(\frac{1}{n}\sum_{1}^{n} Z_j\right)^2 - \left(1 + \frac{1}{n}W_{(1)}(n)\right)^2\right|\right.$$

$$\left. > \frac{L\log n + x}{16n}\right\} + P\left\{\left|\frac{W_{(1)}^2(n)}{n^2}\right| > \frac{L\log n + x}{16n}\right\}$$

$$\leq P_{221} + P_{222} , \quad (5.7)$$

where

$$P_{222} = P\left\{|N(0,1)| > \left(\frac{L\log n + x}{16}\right)^{1/2}\right\}$$

$$\leq \frac{2}{\sqrt{2\pi}} \quad \frac{4}{(L \log n + x)^{1/2}} \quad e^{-(L \log n + x)/32}$$

$$\leq \frac{8 \qquad e^{-x/32}}{\sqrt{2\pi}\ (L \log n + x)^{1/2} n^{L/32}} \tag{5.8}$$

On the other hand,

$$P_{221} = P\left\{ \left| \left[\left(\frac{1}{n} \sum^{n} Z_j\right)^2 - \left(1 + \frac{1}{n} W_{(1)}(n)\right) + \left(1 + \frac{1}{n} W_{(1)}(n)\right) \right]^2 \right.\right.$$

$$\left.\left. - \left(1 + \frac{1}{n} W_{(1)}(n)\right)^2 \right| > \frac{L \log n + x}{16n} \right\}$$

$$= P\left\{ \left| \left[\frac{1}{n} \sum^{n} Z_j - \left(1 + \frac{1}{n} W_{(1)}(n)\right) \right]^2 \right.\right.$$

$$\left. + 2\left[1 + \frac{1}{n} W_{(1)}(n)\right] \left[\frac{1}{n} \sum_{1}^{n} Z_j - \left(1 + \frac{1}{n} W_{(1)}(n)\right)\right] \right|$$

$$\left. > \frac{L \log n + x}{16n} \right\} \leq P\left\{ \left| \frac{1}{n} \sum^{n} Z_j - \left(1 + \frac{1}{n} W_{(1)}(n)\right) \right| \right.$$

$$\left. > \left(\frac{L \log n + x}{32n}\right)^{1/2} \right\} + P\left\{ \left|1 + \frac{1}{n} W_{(1)}(n)\right| \right.$$

$$\left|\frac{1}{n}\sum^{n} Z_j - (1 + \frac{1}{n} W_{(1)}(n))\right| > \frac{L \log n + x}{64n}\Bigg\}$$

$$\leq R_1 + R_2 \tag{5.9}$$

As to R_1 of (5.9), we have

$$R_1 = P\left\{\left|\frac{1}{n}\sum_1^n Z_j - (1 + \frac{1}{n} W_{(1)}(n))\right| > \left(\frac{L \log n + x}{32n}\right)^{1/2}\right\}$$

$$\leq P\left\{\left|\frac{1}{n}\sum_1^n Z_j - (1 + \frac{1}{n} W_{(1)}(n))\right| > \frac{L \log n + x}{32n}\right\} \tag{5.10}$$

provided that $\frac{L \log n + x}{32n} \leq 1$ or $x \leq 32n - L \log n$.

As to R_2 of (5.9), we have for any $\varepsilon > 0$,

$$R_2 \leq P\left\{\left(\left|1 + \frac{1}{n} W_{(1)}(n)\right| \left|\frac{1}{n}\sum_1^n Z_j - (1 + \frac{1}{n} W_{(1)}(n))\right| > \frac{L \log n + x}{64n}\right)\right.$$

$$\left.\cap \left(\left|1 + \frac{1}{n} W_{(1)}(n)\right| \leq 1 + \varepsilon\right)\right\}$$

$$+ P\left\{\left|1 + \frac{1}{n} W_{(1)}(n)\right| > 1 + \varepsilon\right\}$$

$$\leq P\left\{ \left|\frac{1}{n}\sum_{1}^{n} Z_j - (1 + \frac{1}{n} W_{(1)}(n))\right| > \frac{L \log n + x}{64(1+\varepsilon)n} \right\}$$

$$+ P\left\{ |\dot{N}(0,1)| > \varepsilon n^{1/2} \right\}$$

$$\leq R_{21} + R_{22} \tag{5.11}$$

where

$$R_{22} \leq \frac{2}{\sqrt{2\pi}} \frac{1}{\varepsilon n^{1/2}} e^{-\frac{\varepsilon^2 n}{2}} \tag{5.12}$$

Now, with A, B, C as in (5.2), we get for all $x \leq 32n - L \log n$ and with L large enough that

$$P_{21} \leq B\, e^{-Cx/4} ,$$

$$P_{221} \leq R_1 + R_2 \leq B\, e^{-Cx/32} + B\, e^{-Cx/64(1+\varepsilon)} + \sqrt{2/\pi}\, \frac{1}{\varepsilon n^{1/2}}\, e^{-\frac{\varepsilon^2 n}{2}} . \tag{5.13}$$

Hence

$$P_2 \leq B\, e^{-Cx/4} + \frac{8}{\sqrt{2\pi}} \frac{e^{-x/32}}{(L \log n + x)^{1/2} n^{L/32}}$$

$$+ B\, e^{-Cx/32} + B\, e^{-Cx/64(1+\varepsilon)} + \sqrt{2/\pi}\, \frac{1}{\varepsilon n^{1/2}}\, e^{-\frac{\varepsilon^2 n}{2}} . \tag{5.14}$$

In order to estimate P_1 of (5.3), we argue as follows;

$$P_1 = P\left\{ \left| \frac{1}{(\frac{1}{n}\sum_1^n Z_j)^2} - 1 \right| \sup_{0<u\le 1} \left| \frac{1}{n}\sum_1^{[nu]} Z_j^2 - 2u(\frac{1}{n}\sum_1^n Z_j)^2 \right| > \frac{L \log n + x}{2n} \right\}$$

$$\le P\left\{ \left| \frac{1}{(\frac{1}{n}\sum_1^n Z_j)^2} - 1 \right| > \frac{(L \log n + x)^{1/2}}{n} \right\}$$

$$+ P\left\{ \sup_{0<u<1} \left| \frac{1}{n}\sum_1^{[nu]} Z_j^2 - 2u(\frac{1}{n}\sum_1^n Z_j)^2 \right| > \frac{1}{2}(L \log n + x)^{1/2} \right\}$$

$$\le P_{11} + P_{12} \tag{5.15}$$

Now, since

$$\left| \frac{1}{(\frac{1}{n}\sum Z_j)^2} - 1 \right| \le \left| \frac{1}{(\frac{1}{n}\sum Z_j)} - 1 \right|^2 + 2\left| \frac{1}{(\frac{1}{n}\sum Z_j)} - 1 \right| ,$$

then

$$P_{11} \le P\left\{ \left| \frac{1}{(\frac{1}{n}\sum Z_j)} - 1 \right| > \frac{(L \log n + x)^{1/4}}{\sqrt{2}\, n^{1/2}} \right\}$$

$$+ P\left\{ \left| \frac{1}{(\frac{1}{n}\sum^{n} Z_j)} - 1 \right| > \frac{(L \log n + x)^{1/2}}{4n} \right\} \tag{5.16}$$

On the other hand, since

$$\sup_{0<u\le 1} \left| \frac{1}{n} \sum_{1}^{[nu]} Z_j^2 - 2u(\frac{1}{n} \sum_{1}^{n} Z_j)^2 \right|$$

$$\le \sup_{0<u\le 1} \left| \frac{1}{n} \sum_{1}^{[nu]} (Z_j^2 - 2) \right| + 2 \left| (\frac{1}{n} \sum_{1}^{n} Z_j)^2 - 1 \right|$$

$$\le \sup_{0<u\le 1} \left| \frac{1}{n} \sum_{1}^{[nu]} (Z_j^2 - 2) \right| + 2 \left| (\frac{1}{n} \sum_{1}^{n} Z_j) - 1 \right|^2$$

$$+ 4 \left| (\frac{1}{n} \sum_{1}^{n} Z_j) - 1 \right| ,$$

then

$$P_{12} \le P\left\{ \sup_{0<u\le 1} \left| \frac{1}{n} \sum_{1}^{[nu]} (Z_j^2 - 2) \right| > \frac{(L \log n + x)^{1/2}}{4} \right\}$$

$$+ P\left\{ \left|\frac{1}{n} \sum_{1}^{n} Z_j - 1\right| > \frac{(L \log n + x)^{1/4}}{4} \right\}$$

$$+ P\left\{ \left|\frac{1}{n} \sum_{1}^{n} Z_j - 1\right| > \frac{(L \log n + x)^{1/2}}{32} \right\} . \tag{5.17}$$

By applying Chernoff's (1952) Theorem to the right hand side of (5.16) and (5.17), we get

$$P_{11} \leq D^* e^{-\delta^* x} \quad \text{and} \quad P_{12} \leq D^* e^{-\delta^* x} , \tag{5.18}$$

where D^* and δ^* are appropriate absolute positive constants. Combining (5.14), (5.15) and (5.18), and by choosing L, D, δ appropriately, we get

$$P \leq D e^{-\delta x} + \frac{8}{\sqrt{2\pi}} \cdot \frac{e^{-x/32}}{(L \log n + x)^{1/2} n^{L/32}} + \sqrt{2/\pi} \frac{1}{\varepsilon n^{1/2}} e^{-\frac{1}{2}\varepsilon^2 n} . \tag{5.19}$$

This proves the required result.

Define

$$I_n(u) = \begin{cases} 0 & \text{if } u < \frac{2}{n} \\ n^{1/2} \left(n \sum_{1}^{[nu]-1} T_j^2 - 2u \right) & \text{if } \frac{2}{n} \leq u \leq 1 , \end{cases} \tag{5.20}$$

and

$$E_n(u) = \begin{cases} 0 & \text{if } u < \frac{2}{n} \\ n^{1/2} \left[\dfrac{(\frac{1}{n} \sum_1^{[nu]-1} Z_j^2)}{(\frac{1}{n} \sum_1^{n} Z_j)^2} - 2u \right] & \text{if } \frac{2}{n} \leq u \leq 1 \end{cases} \tag{5.21}$$

By Theorem (2.1) and Remark (2.1), we have

$$I_n(u) \xrightarrow{D} K(u;2) \stackrel{D}{=} 2W(u) + (-2 \pm 2\sqrt{5}) \cdot B(u) \quad , \tag{5.22}$$

where $K(u;2)$ is as in (2.17), $W(u)$ is a standard Wiener process and $B(u) = W(u) - uW(1)$ is a Brownian bridge. Then in the light of (5.22) we also have

$$\Psi(I_n(u)) \xrightarrow{D} \Psi(2W(u) + (-2 \pm 2\sqrt{5}) B(u)) \tag{5.23}$$

as $n \to \infty$, for every continuous functional Ψ on the space of real valued functions on $[0,1]$ endowed with the supremum topology. In particular, we have that

$$t_1(I,n) = I_n(1) \xrightarrow{D} 2W(1) \quad , \tag{5.24}$$

$$t_2(I,n) = \sup_{0<u<1} |I_n(u)|$$

$$\xrightarrow{D} \sup_{0<u<1} |2W(u) + (-2 \pm 2\sqrt{5}) B(u)| \quad , \tag{5.25}$$

and

$$t_3(I,n) = \int_0^1 I_n^2(u)du \xrightarrow{D} \int_0^1 (2W(u) + (-2\pm 2\sqrt{5})\ B(u))^2 du \quad . \tag{5.26}$$

In addition to (5.24), (5.25) and (5.26), from Theorem (5.1) we can also prove rates of convergence results for these convergence in distribution. Let $V_{n,i}(x)$ be the distribution function of $t_i(I,n)$, $i = 1, 2, 3$ and let

$$\left.\begin{aligned} V_1(x) &= P\ \{2W(1) \le x\} \\ V_2(x) &= P\ \{\sup_{0<u<1} |2W(u) + (-2\pm 2\sqrt{5})\ B(u)| \le x\} \\ V_3(x) &= P\ \{\int_0^1 (2W(u) + (-2\pm 2\sqrt{5})\ B(u))^2 du \le x\} \quad . \end{aligned}\right\} \tag{5.27}$$

Then (5.24), (5.25) and (5.26) read, respectively

$$\lim_{n\to\infty} V_{n,i}(x) = V_i(x) \quad , \ i = 1,2,3 \quad . \tag{5.28}$$

Put

$$\Delta_{n,i} = \sup_x |V_{n,i}(x) - V_i(x)| \quad , \ i = 1,2,3 \quad . \tag{5.29}$$

The following three Lemmas are going to be used in proving Theorem (5.2).

Lemma 5.1

$V_2(.)$ is differentiable and

$$\sup_{0<y<\infty} V_2'(y) < +\infty \tag{5.30}$$

Lemma 5.2

$V_3(.)$ is arbitrary many times differentiable and for an arbitrary integer p, $V_3^{(p)}(x) \to 0$ as $x \to \infty$.

Lemma 5.3

For any real $p \geq 0$ and integers $q = 0, 1, 2, \ldots$ the function $x^p V_3^{(q)}(x)$ is bounded on $(0,\infty)$.

Remark 5.1

$V_3^{(a)}$ stands for the ath derivative of V_3, and $v_3^{(a)}$ stands for the ath derivative of the density function $v_3 = V_3'$.

Remark 5.2

The proofs of Lemmas (5.1)-(5.3) will not be given in the present thesis. They will be the subject of a further research.

Theorem 5.2

$$\Delta_{n,i} = O(n^{-1/2}\log n) \quad , \; i = 1,2,3 \; .$$

Proof

First, we prove that

$$\Delta_{n,1} = O(n^{-1/2}\log n) \tag{5.31}$$

By (2.5) we have

$$\begin{aligned} V_{n,1}(x) &= P\{I_n(1) \le x\} \\ &= P\{E_n(1) \le x\} \\ &= P\{K_n(1;2) \le x - (E_n(1) - K_n(1;2))\} \\ &\le P\{K_n(1;2) \le x + |E_n(1) - K_n(1,2)|\} \; . \end{aligned}$$

Now, by Theorem (5.1) after taking x of Theorem (5.1) to be $\frac{1}{2\delta}\log n$ and writing $\varepsilon_n = (L + \frac{1}{2\delta})\frac{\log n}{n^{1/2}}$, we have for any $\varepsilon > 0$

$$\begin{aligned} V_{n,1}(x) \le\; & P\{K_n(1;2) \le x + \varepsilon_n\} + D\, n^{-1/2} \\ & + \frac{8}{\sqrt{2\pi}\,(L + \frac{1}{2\delta})^{1/2}\,(\log n)^{1/2}\, n^{L/32 + \frac{1}{64\delta}}} \\ & + \sqrt{2/\pi}\,\frac{1}{\varepsilon\, n^{1/2}}\, e^{-\frac{1}{2}\varepsilon^2 n} \end{aligned}$$

$$\leq P\{2W(1) \leq x + \varepsilon_n\} + D\, n^{-1/2} + e_n \quad , \tag{5.32}$$

where

$$e_n = \frac{8}{\{2\pi \quad (L + \frac{1}{2\delta})\}^{1/2} \; (\log n)^{1/2} \; n^{\frac{2\delta L+1}{64\delta}}}$$

$$+ \sqrt{2/\pi}\, \frac{1}{\varepsilon n^{1/2}}\, e^{-\frac{1}{2}\varepsilon^2 n} \quad . \tag{5.33}$$

Hence

$$V_{n,1}(x) \leq V_1(x + \varepsilon_n) + D\, n^{-1/2} + e_n \quad . \tag{5.34}$$

Similarly, we have

$$V_{n,1}(x) \geq V_1(x - \varepsilon_n) - D\, n^{-1/2} - e_n \quad . \tag{5.35}$$

Therefore

$$\Delta_{n,1} \leq \sup_x P\{x - \varepsilon_n \leq 2W(1) \leq x + \varepsilon_n\} + D\, n^{-1/2} + e_n$$

$$\leq \sup_x \int_{x-\varepsilon_n}^{x+\varepsilon_n} \frac{1}{2\sqrt{2\pi}}\, e^{-t^2/8}\, dt + D\, n^{-1/2} + e_n$$

$$\leq \frac{\varepsilon_n}{\sqrt{2\pi}} + D\, n^{-1/2} + e_n \quad . \tag{5.36}$$

Whence

$$\Delta_{n,1} = O(n^{-1/2}\log n)$$

as claimed.

Second, we prove that

$$\Delta_{n,2} = O(n^{-1/2}\log n) \quad . \tag{5.37}$$

Again by (2.5) we have

$$\begin{aligned} V_{n,2}(x) &= P\ \{\sup_{0<u<1} |I_n(u)| \leq x\} \\ &= P\ \{\sup_{0<u<1} |E_n(u)| \leq x\} \\ &= P\ \{\sup_{0<u<1} |(E_n(u) - K_n(u;2)) + K_n(u;2)| \leq x\} \\ &\leq P\ \{\sup_{0<u<1} |K_n(u;2)| \leq x - \varepsilon_n\} - D\ n^{-1/2} - e_n\ , \end{aligned} \tag{5.38}$$

where the last inequality follows by Theorem (5.1) and e_n , ε_n are as in (5.32) and (5.33). Now

$$\begin{aligned} V_{n,2}(x) \geq P\ \{\sup_{0<u<1} |2W(u) + (-2-2\sqrt{5})\ B(u)| \leq x - \varepsilon_n \\ - D\ n^{-1/2} - e_n \end{aligned}$$

$$\geq V_2(x - \varepsilon_n) - D\, n^{-1/2} - e_n \quad .$$

Similarly we can prove that

$$V_{n,2}(x) \leq V_2(x + \varepsilon_n) + D\, n^{-1/2} + e_n \quad .$$

Hence

$$\Delta_{n,2} \leq \sup_x P\,\{x - \varepsilon_n \leq \sup_{0<u<1} |2W(u) - 2(1+\sqrt{5})\, B(u)|$$

$$\leq x + \varepsilon_n\} + D\, n^{-1/2} + e_n$$

$$\leq \sup_x \int_{x-\varepsilon_n}^{x+\varepsilon_n} v_2(t)\, dt + D\, n^{-1/2} + e_n$$

$$\leq 2\, \varepsilon_n \sup_{0<y<\infty} v_2(y) + D\, n^{-1/2} + e_n$$

$$\leq C\, \varepsilon_n \quad ,$$

where the positive absolute constant C exists, since $\sup_{0<y<\infty} v_2(y)$ is finite by Lemma (5.1). Hence the proof of (5.37) is complete.

Third, we prove that

$$\Delta_{n,3} = O(n^{-1/2} \log n) \quad . \tag{5.39}$$

The basic idea of the following proof is due to J.H. Venter (cf. Remark 3 of S. Csörgő (1980) and Theorem 1, Cotterill and M. Csörgő (1980)).

Let $\omega_n = (t_3(I,n))^{1/2} = (\int_0^1 I_n^2(u)\,du)^{1/2}$ and $\omega_{(n)} = (\int_0^1 K_n^2(y;2)dy)^{1/2}$. Since the latter are norms, we have (this is the idea of J.H. Venter)

$$|\omega_n - \omega_{(n)}| \leq \left(\int_0^1 (I_n(y) - K_n(y;2))^2\,dy\right)^{1/2}$$

$$\leq \sup_{0<y<1} |I_n(y) - K_n(y;2)| = \delta_n$$

Therefore

$$V_{n,3}(x) = P\,\{\omega_n^2 \leq x\} = P\,\{\omega_n \leq x^{1/2}\}$$

$$\leq P\,\{\omega_{(n)} \leq x^{1/2} + \delta_n\}$$

$$\leq P\,\{\omega_{(n)} \leq x^{1/2} + \varepsilon_n\} + D\,n^{-1/2} + e_n$$

$$= P\,\{\omega_{(n)}^2 \leq (x^{1/2} + \varepsilon_n)^2\} + D\,n^{-1/2} + e_n$$

$$= P\,\{\omega^2 \leq (x^{1/2} + \varepsilon_n)^2\} + D\,n^{-1/2} + e_n$$

$$= P\,\{\int_0^1 (2W(u) + (-2-2\sqrt{5})\,B(u))^2\,du \leq (x^{1/2} + \varepsilon_n)^2\} + D\,n^{-1/2} + e_n$$

$$= V_3((x^{1/2} + \varepsilon_n)^2) + D\,n^{-1/2} + e_n \quad , \qquad (5.40)$$

where the third line of (5.40) follows from Theorem (5.1), with ε_n, e_n are as in (5.32) and (5.33). Also, to get the fifth line of (5.40) we have used the following

$$P\{\omega^2_{(n)} \le a\} = P\{\int_0^1 K_n^2(y;2)\,dy \le a\}$$

$$= P\{\int_0^1 K_n^2(y;2)\,dy \le a\}$$

$$= P\{\omega^2 \le a\} \quad ,$$

where

$$\omega^2 = \int_0^1 (2W(y) - 2(1+5)\,B(y))^2\,dy \quad .$$

Similar to (5.40), we can show that

$$V_{n,3}(x) \ge V_3((x^{1/2} - \varepsilon_n)^2) - D\,n^{-1/2} - e_n \quad .$$

Whence

$$\Delta_{n,3} \le \sup_x \{V_3((x^{1/2} + \varepsilon_n)^2) - V_3((x^{1/2} - \varepsilon_n)^2)\}$$

$$+ D\,n^{-1/2} + e_n \quad . \tag{5.41}$$

Now, if $\omega = \{\int_0^1 (2W(u) - (2+2\sqrt{5})\,B(u))^2\,du\}^{1/2}$ and if f is the density function of ω, then $f(y^{1/2}) = 2y^{1/2}\,v_3(y)$, where v_3 is

the density function of ω^2 i.e., $v_3 = V_3'$. Also

$$\Delta_{n,3} \leq \sup_x P\{x^{1/2} - \varepsilon_n \leq \omega \leq x + \varepsilon_n\} + D\, n^{-1/2} + e_n$$

$$\leq \sup_x \int_{x^{1/2}-\varepsilon_n}^{x^{1/2}+\varepsilon_n} y^{1/2}\, v_3(y)\, dy + D\, n^{-1/2} + e_n$$

$$\leq 2\, \varepsilon_n \sup_{0<y<\infty} y^{1/2}\, v_3(y) + D\, n^{-1/2} + e_n$$

$$\leq C_1\, \varepsilon_n \quad ,$$

where the positive absolute constant C_1 exists, since by Lemma (5.3) $y^{1/2}\, v_3(y)$ is bounded on $(0,+\infty)$. Hence the proof of (5.39) is complete. This completes the proof of Theorem (5.2).

The same proof of Theorem (5.2), can be employed to obtain rates of convergence results for the processes $p_n(u;2)$ and $J_n(u;2)$ of (1.4) and (1.7) respectively. Let $V_{n,i}^{(p)}(x)$ be the distribution function of $t_i(p,n)$, $i = 1, 2, 3$, where $t_i(p,n)$ are as in (4.5), (4.6) and (4.7). Also, let $V_{n,i}^{(J)}(x)$ be the distribution function of $t_i(J,n)$, $i = 1, 2, 3$, where $t_i(J,n)$ are as in (4.12), (4.13) and (4.14). Then (4.5), (4.6) and (4.7) read, respectively

$$\lim_{n\to\infty} V_{n,i}^{(p)}(x) = V_i(x) \qquad , \; i = 1,2,3 \quad ,$$

and (4.12), (4.13) and (4.14) read, respectively

$$\lim_{n\to\infty} V_{n,i}^{(J)}(x) = V_i(x) \qquad , \quad i = 1,2,3 \ ,$$

where $V_i(x)$, $i = 1, 2, 3$ are as in (5.27).

Put

$$\Delta_{n,i}(p) = \sup_x |V_{n,i}^{(p)}(x) - V_i(x)| \qquad , \ i = 1,2,3$$

and

$$\Delta_{n,i}(J) = \sup_x |V_{n,i}^{(J)}(x) - V_i(x)| \qquad , \ i = 1,2,3 \ .$$

<u>Theorem (5.3)</u>

(i) $\Delta_{n,i}(p) = O(n^{-1/2}\log n) \qquad , \ i = 1,2,3 \ ,$

(ii) $\Delta_{n,i}(J) = O(n^{-1/2}\log n) \qquad , \ i = 1,2,3 \ .$

CHAPTER V

ON CLUSTERING

1. Introduction, Preliminaries and Definitions

In many applications it is useful to know if a set of observations can be split into two or more groups. Our task is to derive the asymptotic null distribution of a test for homogeneity corresponding to a standard clustering technique which partitions the observations into two groups so that the between-groups sum of squares is maximized.

Let F be a continuous distribution function with unknown location and scale parameters $-\infty < \mu < +\infty$ and $\sigma > 0$ respectively, and assume that F is of the form $F(x;\mu,\sigma) = F_0((x-\mu)/\sigma)$, x real, where $F_0(\ )$ is a known distribution function with mean zero and variance one, i.e., $F \in \mathcal{F}$ of (1.1) of chapter II. Further, assume that F_0 satisfies

$$\left.\begin{array}{l} \int_{-\infty}^{\infty} x \, d F_0 = 0 \quad , \quad \int_{-\infty}^{\infty} x^2 \, d F_0 = 1 \quad , \quad F_0(\tfrac{1}{2}) = 0 \\ \text{and} \quad 2f_0(0)\, m_0 < 1 \text{ , where} \\ \quad m_0 = -2\int_{-\infty}^{0} x \, d F_0 \end{array}\right\} \tag{1.1}$$

Let

$$V_0(1) = 1 - m_0^2 \quad ,$$
$$V_0(2) = \frac{1}{4}\int_{-\infty}^{\infty} x^4 f_0(x)\, dx - \frac{1}{4} \quad . \tag{1.2}$$

Suppose we are given n observations $X_1, X_2, \ldots, X_n$, and we want to check the hypothesis that they are drawn independently from a common distribution function in F, of (1.1) of chapter I, against the alternative that they come from at least two distinct distributions. For any partition $P = (P_1, P_2)$ of the set $\{1, 2, \ldots, n\}$ into two sets, the corresponding between-groups sum of squares is given by

$$B_p = n_1 n_2 (\bar{y}_1 - \bar{y}_2)^2/n \quad , \tag{1.3}$$

where n_i is the number of elements in P_i and $\bar{y}_i = \sum_{P_i} X_j/n_i$ for $i = 1, 2$. Let B denote the maximum value of B_p taken over all partitions P and let $T = \sum_{i=1}^{n} (X_i - \bar{X})^2$. Our task is to derive the asymptotic null distribution of $B/(T/n)$.

Assuming that F_0 is standard normal, estimated percentage points of the null distribution of $B/(T/n)$, obtained by simulation, are given by Engelman and Hartigan (1969) and an empirical expression for the asymptotic distribution is given by Hartigan (1975). Moreover, Brillinger, Knott and Scott (1979) have studied the asymptotic distribution of $B/\hat{\sigma}^2$ with $\hat{\sigma} = (T + \nu s^2)/(n + \nu)$, where s^2 is an independent estimate of σ^2 which is normally available in applications of the AID program and in the analysis of variance. Assuming that $\nu s^2/\sigma^2 \sim \chi^2_\nu$,

for some ν , they have proved that $[(B/\hat{\sigma}^2)^{1/2} - (2n/\pi)^{1/2}]$ is asymptotically normal with mean 0 and variance $1 - w/\pi$ when $\nu/n \to \rho$ as $n \to \infty$, where $w = (3 + 2\rho)/(1 + \rho)$. Applying their result to obtain the asymptotic distribution of $B/(T/n)$, we get that $[(B/(T/n))^{1/2} - (2n/\pi)^{1/2}]$ is asymptotically normal with mean 0 and variance $1 - 3/\pi$.

For general F Hartigan (1978) has shown that the asymptotic distribution of the maximum F-ratio, F max., is normal where F maximum is the maximum value of $B_p/(T_1 + T_2)$, where $T_i = \sum_{p_i} (X_j - \bar{y}_i)^2$ for $i = 1, 2$, taken over all partitions p . In the following we summarize some of Hartigan's (1978) results in terms of our notation.

The quantile function Q for the distribution function F is defined by

$$Q(p) = \sup\{x: F(x) \leq p\} \qquad 0 \leq p \leq 1 .$$

The function Q is nondecreasing and right continuous. If U is uniformly distributed over $[0,1]$, then $Q(U)$ has distribution function F .

Define the lower and upper means of Q at p by

$$\underline{Q}(p) = \frac{1}{p} \int_{q \leq p} Q(p)\, dq \quad , \quad 0 < p \leq 1 ,$$

$$\bar{Q}(p) = \frac{1}{1 - p} \int_{q > p} Q(q)\, dq , \quad 0 < q < 1 .$$

The split function of Q at p is

$$B(Q,p) = p\,\underline{Q}^2(p) + (1-p)\bar{Q}^2(p) - \left(\int_0^1 Q(q)\,dq\right)^2 \quad , \; 0 < p < 1 \; .$$

For the case of two clusters $B(Q,p)$ corresponds to a between groups sums of squares. Define the ratio function.

$$R(Q,p) = B(Q,p)/\{\int_0^1 Q^2(p)\,dp - B(Q,p) - (\int_0^1 Q(p)\,dp)^2\} \; .$$

If $X_1, X_2, \ldots, X_n, \ldots$ is a sequence of independent observations from F and if $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ denote the order statistics of $X_1, \ldots, X_n$ then the empirical random variable Q_n^* is the quantile function of the empirical distribution function F_n of the sample $Q(U_1), \ldots, Q(U_n)$, where $U_i = F(X_i)$, then

$$Q_n^*(p) = Q(U_{i:n}) \qquad \text{for} \quad \frac{i-1}{n} \le p < \frac{1}{n} \; , \quad 1 \le i \le n \; ,$$

where $U_{i:n} = F(X_{i:n})$. (We notice that the definition of Q_n^* is slightly different from the definition of the sample quantile function Q_n which is going to be given in (1.4)).

Let the sample split function $B(Q_n^*,p)$ have its maximum at the sample split point p_n .

<u>Theorem 5A</u> (Hartigan (1978))

Suppose that X has finite fourth moment, and that $B(Q,p)$ has a unique maximum at p_o . Assume that Q has a continuous derivative in the neighbourhood of p_o , and that $[(d^2/dP^2)\,B(Q,P)]_{p=p_o} < 0$. Then

$$n^{1/2}(P_n - P_o) \xrightarrow{D} N(0,\sigma_p^2)$$

$$n^{1/2}(B(Q_n^*,P_n) - B(Q,P_o)) \xrightarrow{D} N(0,\sigma_B^2)$$

$$n^{1/2}(R(Q_n^*,P_n) - R(Q,P_o)) \xrightarrow{D} N(0,\sigma_p^2)$$

where σ_p^2 , σ_B^2 , σ_R^2 are given in (3) and (4) in Hartigan (1978).

Hartigan (1978) has applied the above results to the normal case to get

$$P_n \sim N(0.5, \frac{1}{4n} + \frac{1}{n(2\pi - 4)})$$

$$B(Q_n^*,P_n) \sim N(\frac{2}{\pi}, \frac{8}{\pi}(1 - \frac{2}{\pi})/n)$$

$$R(Q_n^*,P_n) \sim N(\frac{2}{\pi - 2}, \frac{8}{\pi}(1 - \frac{3}{\pi})/(1 - \frac{2}{\pi})^4 n).$$

As a check on the asymptotic formulae, Hartigan (1978) has considered samples of size $n = 10$ and $n = 100$ from the normal distribution. "It is seen that the computed means and variances approximate the theoretical ones. Normal plots reveal that the quantities $B(Q_n^*,P_n)$ and $R(Q_n^*,P_n)$ are quite skew for small n, so that the normal approximation is not too safe for $n < 100$. However $B^{1/2}(Q_n^*,P_n)$ and $R^{-1/2}(Q_n^*,P_n)$ were found by experimentation to be quite normal even for $n = 5$ " as quoted from Hartigan (1978).

To date, as far as we know, the asymptotic distribution of $K = B/(T/n)$ was studied only for the quoted cases, and our own thoughts pick up from here. Our main result is that, under the conditions of our theorem (3.1), $[K^{1/2} - (n\, m_0^2)^{1/2}]$ is asymptotically normal with mean zero and variance $1 - m_0^2\,(1 + V_0(2))$, where m_0, and $V_0(2)$ are defined in (1.2). In particular, if we consider the case of $F_0() = \Phi()$, it is readily seen that the conditions of Theorem (3.1) are satisfied, and consequently $[K^{1/2} - (\frac{2n}{\pi})^{1/2}]$ is asymptotically normal with mean zero and variance $1 - \frac{3}{\pi}$. This result agrees with that of Brillinger, Knott and Scott (1979). Moreover, just like Brillinger, Knott and Scott (1979), we have high hopes that the χ^2 approximation of K, given by corollary (3.2), which suggests approximating k by $c\chi^2_{v_0}$ with $c = 2(1 - m_0^2(1 + V_0(2)))$, and $v_0 = (n\, m_0^2/2c) + \frac{1}{2}$ where $V_0(2)$ is as in (1.2), is going to work very well even for small n.

Suppose $Y_1, Y_2, \ldots, Y_n$ are independent random variables with a continuous distribution function $F(.)$. Let $Y_{1:n} \leq Y_{2:n} \leq \ldots \ldots \leq Y_{n:n}$ denote their order statistics. Define the sample quantile function $Q_n(y)$ as follows:

$$Q_n(y) = Y_{k:n} \quad ; \quad \text{if } \frac{k-1}{n} < y \leq \frac{k}{n} \quad , \quad k = 1,2,\ldots,n. \qquad (1.4)$$

In the sequel it will be always assumed that the underlying distribution $F_0(.)$ satisfies conditions (3.4) (i)-(iii) of chapter II together with conditions (1.6) and (1.9) of Theorem 1A.

2. The Asymptotic Distribution of B

In this section we are interested in the large sample distribution of B ; the maximum between-groups sum of squares. To do that we will consider first the large sample distribution of $M_n = (B/n)^{1/2}$. Let

$$S_j = \sum_{i \le j} \frac{(Y_{i:n} - \bar{y})}{\sqrt{j(n-j)}} \qquad \text{for } j = 1,2,\ldots,n-1 \tag{2.1}$$

Then it is straight forward to verify that

$$M_n = \sup_{1 \le j \le n-1} |S_j| \tag{2.2}$$

and that

$$S_j = \{\frac{1}{n} \sum_{i \le j} Y_{i:n} - \frac{j}{n^2} \sum_{i=1}^{n} Y_{i:n}\} / \sqrt{\frac{j}{n}(1 - \frac{j}{n})} \tag{2.3}$$

Let $B_n(y)$ be the Brownian bridge of Theorem 1A and define

$$g(\theta) = - \int^{\theta} F_0^{-1}(y)\, dy / \theta^{1/2}(1-\theta)^{1/2} \tag{2.4}$$

and

$$Z_n(\theta) = \left[- \int_0^{\theta} \frac{B_n(y)}{f_0(F_0^{-1}(y))} dy + \theta \int_0^1 \frac{B_n(y)}{f_0(F^{-1}(y))} dy \right] / \sqrt{\theta(1-\theta)}, \tag{2.5}$$

for $0 \le \theta \le 1$.

Lemma 5B (Brillinger, Knott and Scott (1979)).

For all $\delta > 0$ and for any $\alpha_n = o(1)$, we have

$$\sup_{\substack{j < n\alpha_n \\ j > n(1-\alpha_n)}} |S_j| \underset{=}{\text{a.s.}} O(\alpha_n^{1/2-\delta}) \tag{2.6}$$

In the following two Lemmas we suppose that $\alpha_n = o(1)$, with $\alpha_n^{-1} = o(\frac{n}{\log n})$ as $n \to \infty$.

Lemma 2.1

Assume that F_0 satisfies all the conditions of Theorem (3.1) of chapter II. Then

$$\sup_{n\alpha_n \le j \le n(1-\alpha_n)} |S_j| = \sup_{n\alpha_n \le j \le n(1-\alpha_n)} [g(j/n) + n^{-1/2} Z_n(j/n)] + e_{1n} \tag{2.7}$$

where

$$e_{1n} \overset{P}{=} o(n^{-1/2} \alpha_n^{-1/2})$$

The proof follows by direct application of Theorem (3.1) of chapter II.

Lemma 5C (Brillinger, Knott and Scott (1979))

$$\sup_{n\alpha_n \le j \le n(1-\alpha_n)} [g(j/n) + n^{-1/2} Z_n(j/n)]$$

$$= \sup_{\alpha_n \le \theta \le 1-\alpha_n} [g(\theta) + n^{-1/2} Z_n(\theta)] + e_{2n} , \tag{2.8}$$

where

$$e_{2n} \underline{\underline{a.s.}} O(n^{-1} \alpha_n^{-3/2} \sqrt{\log n}).$$

Lemma 5D (Brillinger, Knott and Scott (1979))

If the $F_0(\ )$ is the distribution function of a standard normal and if $\hat{\theta}(\alpha_n)$ maximizes $g(\theta) + n^{-1/2} Z_n(\theta)$ over the range $\alpha_n \le \theta \le 1-\alpha_n$, then

$$|\hat{\theta}(\alpha_n) - \frac{1}{2}| \underline{\underline{a.s.}} O((\frac{\log n}{n})^{1/2}) . \tag{2.9}$$

Lemma 2.2

If $F_0(\ )$ satisfies the conditions of Theorem (3.1) of chapter II and if $F_0(\ )$ satisfies (1.1) such that $g_n'(\theta)$ is a.s. continuous and $g''(\theta)$ is continuous in a neighbourhood of $\theta = \frac{1}{2}$. Then

$$|\hat{\theta}(\alpha_n) - \tfrac{1}{2}| \underline{\underline{a.s.}} \; O((\frac{\log n}{n})^{1/2}) \quad . \tag{2.10}$$

Theorem 2.1

$n^{1/2}(M_n - m_0)$ is asymptotically normal with mean zero and variance $1 - m_0^2$ where m_0 is as in (1.1).

Proof

Take $\alpha_n = n^{-\varepsilon}$, where $\varepsilon > 0$ is sufficiently small. Then by (2.6), (2.7) and (2.9) we have

$$M_n = \sup_{1 \le j \le n-1} |S_j| \overset{P}{=} g(\hat{\theta}) + n^{-1/2} Z_n(\hat{\theta}) + o(1) \ ,$$

and it follows from Lemma (2.2) that

$$g(\hat{\theta}) + n^{-1/2} Z_n(\hat{\theta}) = g(\tfrac{1}{2}) + n^{-1/2} Z_n(\tfrac{1}{2}) + e_{3n} \quad ,$$

where

$$e_{3n} \underline{\underline{a.s.}} \; o(n^{-1/2}) \quad .$$

Thus M_n has the same asymptotic distribution as $g(\frac{1}{2}) + n^{-1/2} Z_n(\frac{1}{2})$. But $g(\frac{1}{2}) = m_0$, and $Z_n(\frac{1}{2})$ is a normal random variable with mean zero and variance $1 - m_0^2$. Hence $n^{1/2}(M_n - m_0)$ is asymptotically normal with mean zero and variance

$$E\, Z_n^2(\tfrac{1}{2}) = E\left\{ \int_0^1 \frac{B_n(y)}{f_0(F_0^{-1}(y))}\, dy - 2\int_0^{1/2} \frac{B_n(y)}{f_0(F_0^{-1}(y))}\, dy \right\}^2$$

$$= \int_0^1 \int_0^1 \frac{E\, B_n(y)\, B_n(x)}{f_0(F_0^{-1}(y))\, f_0(F_0^{-1}(x))}\, dydx - 4 \int_{1/2}^1 \frac{1-y}{f_0(F_0^{-1}(y))}$$

$$\int_{1/2}^1 \frac{x}{f_0(F_0^{-1}(x))}\, dxdy$$

$$= 1 - m_0^2 \quad .$$

To fill in the details, one should be aware of the following: $E_0X = 0 < \infty$ implies that $\lim_{x\to\infty} x(1 - F_0(x)) = 0$ and $\lim_{x\to-\infty} xF_0(x) = 0$, $\frac{1}{2}(E_0X)^2 = \int_{-\infty}^{\infty} xf_0(x) \int_{\infty}^{x} yf_0(y)\, dydx$.

3. The Asymptotic Distribution of $B/\frac{T}{n}$

In order to consider the general case, in which we are given n observations $X_1, X_2, \ldots, X_n$ and we want to check the hypothesis that they are drawn independently from a common distribution function $F \in \mathcal{F}$, we first note that under the null hypothesis

$$X_i = \sigma Y_i + \mu \quad , \qquad (i = 1,2,\ldots,n) \ , \tag{3.1}$$

where the Y_i are i.i.d.rv with distribution function F_0 .

Then it is straightforward to verify that

$$n^{1/2}\left(\frac{M_n}{\sigma} - m_0\right) \tag{3.2}$$

is asymptotically normal with mean zero and variance $1 - m_0^2$.

To find the asymptotic distribution of $K = \frac{B}{T/n}$ it is simpler to work with $\left(\frac{K}{n}\right)^{1/2} = M_n/\sqrt{T/n}$. We begin with Lemma (3.4) of Csörgő, M. and Révész, P. (1979).

<u>Lemma 5E</u> (Csörgő and Révész (1979))

If in addition to (1.1) F_0 also satisfies:

$$\lim_{y\to 0} y^{1/r}\,|F_0^{-1}(y)| = \lim_{y\to 1} (1-y)^{1/r}\,F_0^{-1}(y) = 0\,, \tag{3.3}$$

for some $r > 4$, and

$$\inf_{0<y<1} \frac{|f_0'(F_0^{-1}(y))|}{y^{\delta}} > 0\,, \quad \inf_{0<y<1} \frac{|f_0'(F_0^{-1}(y)|}{(1-y)^{\delta}} > 0\,, \tag{3.4}$$

for some $1 \le \delta \le 3/2$, then

$$\frac{\sqrt{\frac{T}{n}} - \sigma}{\sigma} = n^{-1/2}\, I_n^1 + o_p(n^{-1/2})\,, \tag{3.5}$$

where

$$I_n^1 = n^{-1} \sum_{k=1}^{n} \frac{\rho_n^o\left(\frac{k}{n+1}\right)}{f_0\left(F_0^{-1}\left(\frac{k}{n+1}\right)\right)} \left(F_0^{-1}\left(\frac{k}{n+1}\right)\right)\,, \tag{3.6}$$

where $\rho_n^o(.)$ is as in (1.9) of chapter II.

In addition to Lemma 1E, Csörgő and Révész (1979) have pointed out that

$$|I_n^1 - \int_0^1 B_n(y)\, 2^{-1}\, d(F_0^{-1}(y))^2| \overset{P}{\to} 0 . \tag{3.7}$$

A direct, but rather tedious calculation yields that $\int_0^1 B_n(y)\, 2^{-1}\, d(F_0^{-1}(y))^2$ is a normal random variable, with mean zero and variance

$$V_0(2) = \frac{1}{4}\int_{-\infty}^{\infty} x^4\, f_0(x)dx - \frac{1}{2}\int_{-\infty}^{\infty} x^2\, f_0(x) \int_{-\infty}^{\infty} y^2\, f_0(y)dy\, dx$$

$$= \frac{1}{4}\int_{-\infty}^{\infty} x^4\, f_0(x)dx - \frac{1}{2}\,(\tfrac{1}{2}(E_0X^2)^2)$$

$$= \frac{1}{4}\int_{-\infty}^{\infty} x^4\, f_0(x)dx - \frac{1}{4} \tag{3.8}$$

In particular, if $f_0(x)$ is the density function of standard normal r.v., then $\int_0^1 B_n(y)\, 2^{-1}\, d(F_0^{-1}(y))^2$ is a $N(0,1/2)$ r.v.

At this stage, we are able to introduce and prove our main result, namely

<u>Theorem 3.1</u>

If in addition to all the conditions of Lemma 5E and Lemma (2.2),

we also have that $M_n/(T/n)^{1/2}$ and $(T/n)^{1/2}$ are independent, then

$$K^{1/2} - (n\, m_0^2)^{1/2} \tag{3.9}$$

is asymptotically normal with mean zero and variance $1 - m_0^2(1 + V_0(2))$.

Proof

Consider

$$K^{1/2} - (n\, m_0^2)^{1/2} = n^{1/2}[(\frac{k}{n})^{1/2} - m_0] = n^{1/2}[\frac{M_n}{\sqrt{T/n}} - m_0]$$

$$= \frac{\sigma}{\sqrt{T/n}}\{n^{1/2}[\frac{M_n}{\sigma} - m_0] - m_0\, n^{1/2}[\frac{\sqrt{T/n} - \sigma}{\sigma}]\}$$

from which we have

$$n^{1/2}[\frac{M_n}{\sqrt{T/n}} - m_0] + m_0 \frac{\sigma}{\sqrt{T/n}}\, n^{1/2}[\frac{\sqrt{T/n} - \sigma}{\sigma}] =$$

$$= \frac{\sigma}{\sqrt{T/n}}\, n^{1/2}[\frac{M_n}{\sigma} - m_0] \tag{3.10}$$

Since

$$\frac{\sigma}{\sqrt{T/n}} \xrightarrow{P} 1 , \tag{3.11}$$

and by (3.2), $n^{1/2}[\frac{\bar{M}_n}{\sigma} - m_0]$ is asymptotically normal with mean zero and variance $V_0(1) = 1 - m_0^2$. Then the right hand side of (3.10) is asymptotically normal with mean zero and variance $V_0(1)$.

Now, from (3.5) we have

$$n^{1/2}\,[\frac{\sqrt{T/n} - \sigma}{\sigma} = I_n^1 + o_p(1) \quad ,$$

and (3.7) and (3.8) imply that

$$I_n^1 \xrightarrow{D} N(0,V_0(2)) \text{ r.v.}$$

These facts together with (3.11) implies that

$$\frac{\sigma}{\sqrt{T/n}}\, m_0\, n^{1/2}\,[\frac{\sqrt{T/n} - \sigma}{\sigma}] \xrightarrow{D} N(0,m_0^2 V_0(2)) \text{ r.v.} \tag{3.12}$$

By taking the characteristic function of both sides of (3.10), taking into account the assumption that the two r.v.'s of the left hand side of (3.10) are independent, then by proceeding to the limit, we get that the limiting characteristic function of

$$n^{1/2}[\frac{M_n}{\sqrt{T/n}} - m_0] \qquad \text{is}$$

$$\exp - \frac{1}{2}t^2(1 - m_0^2)/\exp - \frac{1}{2}\, t^2\, m_0^2 V_0(2) \quad . \tag{3.13}$$

Therefore $n^{1/2}\,[\frac{M_n}{\sqrt{T/n}} - m_0]$ is asymptotically normal with mean zero and variance $1 - m_0^2(1 + V_0(2))$. This completes the proof.

Remark 3.1

A sufficient condition for $\frac{M_n}{\sqrt{T/n}}$ and $\sqrt{T/n}$ to be independent is that $(\bar{X},T/n)$ are jointly complete sufficient for $F(.)$. Since the distribution of $M_n/\sqrt{T/n}$ is functionally independent of (μ,σ^2), then by Basu's Theorem (1955) we have that $M_n/\sqrt{T/n}$ and T/n are independent random variables.

Corollary 3.1

If $F(.)$ was the distribution function of a $N(\mu,\sigma^2)$ r.v., and consequently $F_0(\ ) = \Phi$, the distribution function of a standard normal r.v., then, by remark (3.1), $M_n/\sqrt{T/n}$ and $\sqrt{T/n}$ are independent. By Theorem (3.1), we have

$$K^{1/2} - (\frac{2n}{\pi})^{1/2} \xrightarrow{D} N(0,1-\frac{3}{\pi}) \text{ r.v.} \tag{3.14}$$

since, for $F_0(x) = \Phi(x)$, we have

$$m_0 = \sqrt{2/\pi} \quad , \quad V_0(1) = 1 - \frac{2}{\pi} \quad \text{and}$$

$$V_0(2) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2} \tag{3.15}$$

Corollary 3.2

A comparison of (3.9), with the familiar result that

$(2\chi_v^2)^{1/2} - (2v-1)^{1/2}$ is asymptotically standard normal suggests approximating K by $c\chi_{v_0}^2$ where

$$c = 2(1 - m_0^2(1 + V_0^2(2))) \quad \text{and} \quad v_0 = (n\, m_0^2/2c) + \frac{1}{2} \quad . \tag{3.16}$$

In particular, if $F_0(x) = \Phi(x)$, then K can be approximated by $c\chi_{v_0}^2$ with

$$c = 2(\pi - 3)/\pi \quad , \quad \text{and} \quad v_0 = \frac{n}{\pi - 3} + \frac{1}{2} \quad .$$

Remark 3.2

One way of interpreting Corollary (3.2) is in terms of robustness. Namely we can say that having assumed normality, our test statistic for clustering is going to be robust with respect to deviations from the latter to the extent the respective general constants c and v_0 of (3.16) will differ from $c = 2(1 - \frac{3}{\pi})$ and $v_0 = \frac{n}{\pi - 3} + \frac{1}{2}$. Consequently, having assumed normality, our test statistic for clustering is going to be robust with respect to deviations from the latter to the extent the respective constants m_0 , of (1.1), and E_0x^4 will differ from $\sqrt{2/\pi}$ and 3 .

Remark 3.3

Although this section is dealing with the asymptotic distribution of $K = B/(T/n)$ which is a special case of $K' = B/\hat{\sigma}$ with

$\hat{\sigma} = \frac{T + \nu s^2}{n + \nu}$, where s^2 is independent estimate of σ^2 such that $\nu s^2/\sigma^2 \sim \chi^2_\nu$, the proof of our Theorem (3.1) can be employed to deduce similar results concerning the asymptotic distribution of K' . However there will be a little change in the proof subject to whatever information we might have about s^2 in the light of the kind of application one would be dealing with.

REFERENCES

[1] Aly, E.A.A. (1979). On the Asymptotic Distribution of a Test for Clustering. Carleton Mathematical series No. 163, July, 1979.

[2] Aly, E.A.A. (1981). On Exponential and Uniform Spacings. Submitted for publication.

[3] Basu, D. (1955). On Statistics Independent of a Complete Sufficient Statistics. Sankhya 15, 377-80.

[4] Bennett, C.A. (1952). Asymptotic Properties of Ideal Linear Estimators. Unpublished Dissertation, Univ. of Michigan.

[5] Brillinger, D.R., Knott, M. and Scott, A.J. (1979). On the Distribution of a Test for Clustering. Paper in manuscript.

[6] Burke, M.D., Csörgő, M., Csörgő, S. and Révész, P. (1979). Approximations of the Empirical Process when Parameters are Estimated. Annals of Probability, 9, 790-810.

[7] Chan, A.H.C. (1977). On the Increments of Multiparameter Gaussian Processes. Ph.D. Thesis, Carleton University.

[8] Chernoff, H., Gastwirth, J.L. and Johns, M.V., Jr., (1967). Asymptotic Distribution of Linear Combinations of Functions of Order Statistics with Application to Estimation. Ann. Math. Statist. 38, 52-72.

[9] Cotterill, D.S. and Csörgő, M. (1979). On the Limiting Distribution of and Critical Values for the Multivariate Cramér-von Mises statistic. Carleton Mathematical lecture note No. 24, August 1979.

[10] Csörgő, M. and Révész, P. (1975). A New Method to Prove Strassen Type Laws of Invariance Principle. II. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 31, 261-269.

[11] Csörgő, M. and Révész, P. (1978). Strong Approximations of the Quantile Process. Annals of Statistics, 6, 882-294.

[12] Csörgő, M. and Révész, P. (1979). Quadratic ... Ity Tests. Carleton Mathematical Series No. 162.

[13] Csörgő, M. and Révész, P. (1980). Quantile Processes and Sums of Weighted Spacings for Composite Goodness-of-Fit I. Carleton Mathematical Lecture Note No. 27-1980.

[14] Csörgő, M. and Révész, P. (1980). Quantile Processes and Sums of Weighted Spacings for Composite Goodness-of-Fit II. Carleton Mathematical Lecture Note No. 27-1980.

[15] Csörgő, M. and Révész, P. (1981). Strong Approximations in Probability and Statistics. Academic Press, New York.

[16] Csörgő, M., Komlós, J., Major, P., Révész, P. and Tusnády, G. (1974). On the Empirical Process when Parameters are Estimated. In Transactions Seventh Prague Conference 1974, Prague: Academia 1977, 87-97.

[17] Csörgő, S. (1980). Limit Behaviour of the Empirical Characteristic Function. Ann. prob., 9, 130-144.

[18] Darling, D.A. (1953). On a Class of Problems Related to the Random Division of an Interval. Ann. Math. Stat., 24, 239-253.

[19] Darling, D.A. (1955). The Cramér-Smirnov Test in the Parametric Case. Annals of Mathematical Statistics, 26, 1-20.

[20] DeWet, T. and Venter, J.H. (1972). Asymptotic Distributions of Certain Tests Criteria of Normality. South African Statistical Journal, 6, 135-149.

[21] DeWet, T. and Venter, J.H. (1973). A Goodness of Fit Test for a Scale Parameter Family of Distributions. South African Statistical Journal, 7, 35-46.

[22] Durbin, J. (1961). Some Methods of Constructing Exact Tests. Biometrika, 48, 41-55.

[23] Durbin, J. (1973a). Weak Convergence of the Sample Distribution Function when Parameters are Estimated. Annals of Statistics, 1, 279-290.

[24] Durbin, J. (1973b). Distribution Theory for Tests Based on the Sample Distribution Function. Regional Conference Series in Applied Mathematics, 9, Philadelphia: Society for Industrial and Applied Mathematics.

[25] Durbin, J. (1976). Kolmogorov-Smirnov Tests when Parameters are Estimated. Empirical Distributions and Processes, eds. P. Gaenssler, and P. Révész, Lecture Notes in Mathematics 566, Berlin: Springer-Verlag, 33-44.

[26] Engelman, L. and Hartigan, J.A. (1969). Percentage Points of a Test for Clusters. J. Amer. Statist. Assoc., 64, 1947-8.

[27] Hartigan, J.A. (1975). Clustering Algorithms. John Wiley & Sons, New York.

[28] Hartigan, J.A. (1978). Asymptotic Distributions for Clustering Criteria. Ann. Statist. 6, 117-131.

[29] Kac, M., Kiefer, J. and Wolfowitz, J. (1955). On Tests of Normality and Other Tests of Goodness-of-Fit Based on Distance Methods. Annals of Mathematical Statistics, 26, 189-211.

[30] Kiefer, J. (1970). Deviations Between the Sample Quantile Process and the Sample D.F. Non-parametric Techniques in Stat. Inference (M.L. Puri, Ed.) 299-319, Cambridge Univ. Press.

[31] Kiefer, J. (1972). Skorohod Embedding of Multivariate R.V.'s and the Sample D.F., Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 24, 1-35.

[32] Kimball, B.F. (1950). On the Asymptotic Distribution of the Sum of Powers of Unit Frequency Differences. Ann. Math. Statist. 21, 263-271.

[33] Komlós, Major, Tusnády (1975). An Approximation of Partial Sums of Independent R.V.'s and the Sample D.F. I. Z. Wahrscheinlichkeitstheorie Verw. Gebiete 32, 111-131.

[34] Komlós, Major, Tusnády (1976). An Approximation of Parital Sums of Independent R.V.'s and the Sample D.F. II. Z. Wahrscheinlichkeitstheorie Verw. Gebiete 33, 33-58.

[35] Koziol, J.A. (1980). A Note on Limiting Distributions for Spacings Statistics. Z. Wahrscheinlichkeitstheorie Verw. Gebiete 51, 55-62.

[36] LeCam, L. (1958). Un théorème sur la division d'un intervalle par des points prés au hasard. Pub. Inst. Statist. Univ. Paris, 7, 7-16.

[37] Lévy, P. (1937). Théorie de l'addition des variables aléatoires. Gauthier-Villars, Paris.

[38] Lévy, P. (1948). Processus stochastique et mouvement brownien. Gauthier-Villars, Paris.

[39] Moran, P.A.P (1947). The Random Division of an Interval. J. Roy. Stat. Soc. Suppl., Vol. 9, pp. 92-98.

[40] Moore, D.S. (1968). An Elementary Proof of Asymptotic Normality of Linear Functions of Order Statistics. Ann. Math. Statist. 39, 263-265.

[41] Neuhaus, G. (1974). Asymptotic Properties of the Cramér-von Mises Statistic when Parameters are Estimated. Proceedings of the Prague Symposium on Asymptotic Statistics 1973, ed. J. Hajek, Prague: Charles University, 2, 257-297.

[42] Neuhaus, G. (1976a). Asymptotic Properties of the Cramér-von Mises Test under Contiguous Alternatives. Journal of Multivariate Analysis, 6, 95-110.

[43] Neuhaus, G. (1976b). Weak Convergence under Contiguous Alternatives of the Empirical Process when Parameters are Estimated: the D_k Approach. Empirical Distributions and Processes, eds. P. Gaenssler, and P. Révész, Lecture Notes in Mathematics 566, Berlin: Springer-Verlag, 68-82.

[44] Parzen, Emanuel (1979a). Nonparametric Statistical Data Modeling. Journal of the American Statistical Association, 74, 105-131.

[45] Parzen, Emanuel (1979b). A Density-Quantile Function Perspective on Robust Estimation. Robustness in Statistics, eds. Robert L. Launer, and Graham N. Wilkinson, New York: Academic Press, 237-258.

[46] Pitman, E.J.G. (1939). The Estimation of the Location and Scale Parameters of a Continuous Population of Any Given Form. Biometrika 30, 391-421.

[47] Pyke, R. (1965). Spacings. J. Roy. Stat., Ser. B., Vol. 7 pp. 395-449.

[48] Pyke, R. (1972). Spacings Revisited. Proceedings of the sixth Berkeley Symposium on Mathematical Statistics and Probability, I, 417-427.

[49] Rao, K.C. (1972). The Kolmogorov, Cramér-von Mises Chi-square Statistics for Goodness-of-Fit Tests in the Nonparametric Case. Abstract: IMS Bulletin, 1, 87.

[50] Sethuraman, J. and Rao, J.S. (1970). Pitman Efficiencies of Tests Based on Spacings. Nonparametric Techniques in Statistical Inference, (M.L. Puri ed.). Cambridge Univ. Press., 405-416.

[51] Shapiro, S.S. and Francia, R.S. (1972). An Approximate Analysis of Variance Test for Normality. Journal of the American Statistical Association, 67, 215-216.

[52] Shapiro, S.S. and Wilk, M.B. (1965). An Analysis of Variance Test for Normality (complete samples). Biometrika, 52, 591-611.

[53] Stigler, S.M. (1974). Linear Functions of Order Statistics with Smooth Weight Functions. Ann. Statist., 2, No. 4, 676-693.

[54] Stigler, S.M. (1979). Correction to Linear Functions of Order Statistics with Smooth Weight Functions. Ann. Statist., 7, 466.

[55] Weiss, L. (1957). Tests of Fit in the Presence of Nuisance Location and Scale Parameters. Ann. Math. Stat., 28, 1016-1020.

[56] Weiss, L. (1961). On the Estimation of Scale Parameters. Naval Research Logistics Quarterly, 8, 245-256.

[57] Weiss, L. (1963). On the Asymptotic Distribution of an Estimate of a Scale Parameter. Naval Research Logistics Quarterly, 10, 1-9.

END

2 7 H 0 8 H 8 2

FIN