## INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

**The quality of this reproduction is dependent upon the quality of the copy submitted.** Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road. Ann Arbor. MI 48106-1346 USA
313/761-4700 800/521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Order Number 9130290

# On the asymptotic theory of the bootstrap

Arcones, Miguel Angel, Ph.D.

City University of New York, 1991

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

A

# ON THE ASYMPTOTIC THEORY OF THE BOOTSTRAP

by

Miguel A. Arcones

A disertation submitted to the Graduate Faculty
in Mathematics in partial fulfillment of the
requirements of the degree of Doctor of Philosophy,
The City University of New York.

1991

This manuscript has been read and accepted for the Graduate Faculty in Mathematics in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Date

Chair of Examining Committee

5/2/91

Date

Executive Officer

Edgar A. Feldman

Evarist Giné

Jay S. Rosen

Richard S. Sacksteder

Supervisory Committee

The City University of New York

# ABSTRACT

# ON THE ASYMPTOTIC THEORY OF THE BOOTSTRAP

by

Miguel A. Arcones

Adviser: Professor Evarist Giné

This dissertation is a compilation of five papers by the author of this thesis and his advisor about the bootstrap of the central limit theorem in different situations.

The first chapter is an introduction to the topic.

The second chapter continues the work of Giné and Zinn (1990) on the bootstrap of the mean in the absence of second moment. The case of finite second moment was studied by Bickel and Freedman (1981). Giné and Zinn showed that if the bootstrap holds in probability then X must be in the domain of attraction of a normal law and if it holds a.s. then $EX^2 < \infty$. Athreya showed that if X is in the domain of atraction of any stable law, then bootstrap CLT holds in probability if the bootstrap sample size is taken to be much smaller than n, namely if $m_n/n \to 0$. Here we show among other results, that

1) if X is in the domain of attration of a normal random variable the bootstrap always holds in probability and that

2) if X is in a domain of atraction of a stable law and $EX^2 = \infty$, then bootstrap CLT holds a.s. if $m_n \log\log n/n \to 0$ and does not hold a.s. if sup $m_n \log\log n/n > 0$.

3) the convergence of boostrap moments.

In the chapter III we examine the bootstrap of U and V statistics particulary in the degenerate case. The nondegenerate case was studied by Bickel and Freedman. Regarding to the degenerate

case, Bretagnolle showed that the naive bootstrap also holds a.s. for U and V statistics under some additional conditions on the moments if $m_n/\log n \to 0$. Here we present a modified bootstrap CLT that holds a.s. for any $m_n$ in particular for $m_n = n$ under weak integrabiblity conditions. The modification consists in bootstrapping the first term of the Hoeffding expansion with respect to $P_n$. Some applications to testing are given.

The chapter IV deals with the bootstrap of a test of symmetry. We give mathematical justification to the method used in Barker and Schuster (1988). Here. the data must be modified before bootstrapping and in this sense the situation is not standard. Our tests would be a particular case of a general class of tests of Romano (1988), if the centering parameters had some good differentiability properties, but they do not. Moreover we give 3 other different tests.

Finally chapter V contains the bootstrap of some M-estatistics whose asymptotics can be deduced from the theory of empirical measures, through the results about the bootstrap of empirical measures in Giné-Zinn (1990a). The M-estimators we consider are in the framework of Huber (1967) and Pollard (1985). They cover most M-estimators used in practice. We obtain, as examples of the general theory, the a.s. bootstrap of the spatial medians, k-means and Huber estimators.

# ACKNOWLEDGEMENT

Although this dissertation was done from the Fall 1988 to the Summer 1990 in CUNY, I was introduced to Probability Theory during the year before, that I spent at Texas A & M University. I thank the Department of Mathematics at Texas A & M University for its finantial support during one year and for offering courses that introduced me in this topic, specially one by J. Zinn.

I thank also the Department of Mathematics of the Graduate School of CUNY for its finantial suppport and for the oportunity to be exposed to different areas of mathematics. I point out the courses by E. Feldman, E. Giné and J. Rosen.

I thank specially to Professor Evarist Giné, Chair of Committee, for his help in miscellaneous issues, for suggesting topics to work on, and for the creative conversations that are the heart of this work.

Finally, I have to mention that my work in this dissertation has been partly supported by a scholarship of the Joint US-Spain Committee and by NSF and PSC-CUNY grants.

# TABLE OF CONTENTS

page

CHAPTER

I. INTRODUCTION TO THE BOOTSTRAP CENTRAL LIMIT THEOREM

II. THE BOOTSTRAP OF THE MEAN WITH ARBITRARY BOOTSTRAP SAMPLE SIZE

III. THE BOOTSTRAP OF U-STATISTICS AND V-STATISTICS

IV. SOME BOOTSTRAP TESTS OF SYMMETRY FOR UNIVARIATE CONTINUOUS DISTRIBUTIONS

V. THE BOOTSTRAP OF M-ESTIMATORS

# CHAPTER I

# INTRODUCTION TO THE BOOTSTRAP CENTRAL LIMIT THEOREM

## 1. A framework for the bootstrap.

In his landmark 1979 paper, Efron introduced a resampling method for the approximation of the distribution of statistical functionals, the bootstrap. Since then this method has been widely applied and thoroughly studied. This thesis is a contribution to the study of the bootstrap in situations which are not completely standard. Most of the results have been obtained in collaboration with Evarist Giné and constitute five articles, one of them already published. We obtain a better understanding of the asymptotic theory of the bootstrap than our predeccesors did in each problem we treat.

Let us start by introducing this statistical method. Suppose that we have an estimator $T_n(X_1,\dots,X_n)$. The bootstrap method consists of, given a sample $X_1,\dots,X_n$ , take $X^{\omega}_{n,1},\dots,X^{\omega}_{n,n}$ i.i.d. r.v. with law $\mathbf{P}_n(\omega)=n^{-1}\sum_{i=1}^{n}\delta_{X_i}$ and then find $T_n(X^{\omega}_{n,1},\dots,X^{\omega}_{n,n})$ as an approximation of $T_n(X_1,\dots,X_n)$. We call $X^{\omega}_{n,1}\dots,X^{\omega}_{n,n}$ the bootstrap sample and $T_n(X^{\omega}_{n,1},\dots,X^{\omega}_{n,n})$ the bootstrap estimator.

Next, I will show a way to construct the bootstrap samples. Suppose $X_n:(\Omega,\mathcal{A},\mathbf{P})\to(\mathrm{S},\mathcal{S})$ are independent identically distributed (i.i.d.) random variables (r.v.) where S is a metric space with the Borel $\sigma$-algebra. Consider also $([0,1],\mathcal{B}\ ,\lambda)$ and $\xi_i:[0,1]\to[0,1]$ $(\xi_i)$ is a sequence of i.i.d. r.v. with law $U[0,1]$. Then for each $n,m,\omega$ we can obtain $X^{\omega}_{n,1}\dots,X^{\omega}_{n,m}$ with law $\mathbf{P}_n=n^{-1}\sum_{i=1}^{n}\delta_{X_i}$ in the following way: for each $t\in[0,1]$ define $X^{\omega}_{n,i}(t)=X_k$ if $\xi_i(t)\in[(k-1)/n,k/n)$ for $k=1,\dots,n$.

Suppose that we have an estimator $T_n$ so that $T_n(X_1,\dots,X_n)\xrightarrow{\mathcal{L}}T_0$ weakly in S. We expect that the distribution of the bootstrap estimator, conditionally on $\omega$, is close (in an appropiate sense) to the distribution of the original estimator. We are going to make sense of $T_n(X^{\omega}_{n,1},\dots,X^{\omega}_{n,n}) \xrightarrow{\mathcal{L}} T_0$ in two ways. Take any distance $D$ that metrizes the weak topology in S. Conditionally on $\omega$ $\mathcal{L}\,(T_n(X^{\omega}_{n,1},\dots,X^{\omega}_{n,n}))$ is a law in S. Then $D(\mathcal{L}\,(T_n(X^{\omega}_{n,1},\dots,X^{\omega}_{n,n}),\mathcal{L}\,(T_0))$ is a function from $\Omega$ to $\mathbf{R}$. Note that maybe it is not measurable. Given a function $f:\Omega\to\mathbf{R}$ we call $f^*$ the upper measurable cover of f. If

(1.1) $D^*(\hat{\mathcal{L}}\,(T_n(X^{\omega}_{n,1},\dots,X^{\omega}_{n,n}),\mathcal{L}\,(T_0)) \to 0$ a.s. then we talk of a.u. (almost uniformly) bootstrap.

By proposition 1.1 in Dudley (1985) (1.1) is equivalent to the fact that for any $\epsilon$

$\mathbf{P}^*\{\ \sup_{n\geq m} D(\mathcal{L}\,(T_n(X^{\omega}_{n,1},\dots,X^{\omega}_{n,n}),\mathcal{L}\,(T_0)) > \epsilon\} \to 0.$

Note that this definition makes sense because it is independent of the metric we take. If $D_1$ and $D_2$ are two metrics for convergence in law in S, then for each $X$ and $\epsilon > 0$ there is a $\delta > 0$ so that if $D_1(\mathcal{L}\,(X),\mathcal{L}\,(Y)) \leq \delta$ then $D_2(\mathcal{L}\,(X),\mathcal{L}\,(Y)) \leq \epsilon$. Therefore

$\mathbf{P}^*\{\ \sup_{n\geq m} D_2(T_n(X^{\omega}_{n,1},\dots,X^{\omega}_{n,n}),T_0) > \epsilon\} \leq \mathbf{P}^*\{\ \sup_{n\geq m} D_1(\ (T_n(X^{\omega}_{n,1},\dots,X^{\omega}_{n,n}),T_0)\ ) > \delta\}.$

If the r.v.'s $X_i$ take values in a a separable metric space (S,d) then there is no measurability problem because there is a countable class of bounded continuous functions $\mathcal{F}$ so that

$\|\mathbf{P}-\mathbf{Q}\|_{\mathcal{F}} := \sup_{f\in\mathcal{F}} |\mathbf{P}\,f - \mathbf{Q}\,f|$ is a metric for the weak convergence. Then

$\hat{D}^*(\hat{\mathcal{L}}\,(T_n(X^{\omega}_{n,1},\dots,X^{\omega}_{n,n}),\mathcal{L}\,(T_0)) = \sup_{f\in\mathcal{F}} |n^{-n}\sum_{i_1,\dots,i_n=1}^{n} f(T_n(X_{i_1},\dots,X_{i_n})) - \mathbf{E}f(T_0)|$ is measurable.

Let us see that such a collection of functions exist. Take a countable dense set C in S. For each $x_1,\dots,x_m \in C$ $p_1,\dots,p_m \in Q^+$ , $q \leq 1$, and $q \in \mathbf{Q}^+$ define $O = \cup_{j=1}^{m} B(x_j,p_j)$ and $f(x) = (q - d(x,O))^+$. Call $\mathcal{F}$ the class of all functions formed in this way. Since $\mathcal{F} \subset BL_1$ where $BL_1 = \{f : f : S \to \mathbf{R}, \sup_t |f(t)| \leq 1$ and $\sup_{s,t} |f(s) - f(t)| \leq d(t,s)\}$

$\mathbf{P}_n \xrightarrow{\mathcal{L}} \mathbf{P}$ implies $\|\mathbf{P}_n - \mathbf{P}\|_{\mathcal{F}} \to 0$.

Conversely supose $\|\mathbf{P}_n - \mathbf{P}\|_{\mathcal{F}} \to 0$. Take an open set G in S. Order $C \cap G = \{x_j\}$. If

$\epsilon_j :=\sup\{\epsilon : B(x_j,\epsilon) \subset G\}$ if $x_j \in G$ and $\epsilon_j := 0$ if $x_j \notin G$ then $G = \cup_{j=1}^{\infty} B(x_j,\epsilon_j)$. We can take m and $0 < p_j < \epsilon_j$ $p_j \in Q^+$ so that $\mathbf{P}\{G\} \cong \mathbf{P}\{\cup_{j=1}^{m} B(x_j,p_j)\}$. Take $q \in \mathbf{Q}^+$ with $0 < q < 1$ and $p_j + q < \epsilon_j$. Define $O = \cup_{j=1}^{m} B(x_j,p_j)$ and $f(x) = (q - d(x,O))^+$. Then $\liminf \mathbf{P}_n\{G\} \geq \liminf \mathbf{P}_n(f/q) = \mathbf{P}(f/q) \geq \mathbf{P}\{O\}$. So $\liminf \mathbf{P}_n\{G\} \geq \mathbf{P}\{G\}$. Therefore $\mathbf{P}_n \xrightarrow{\mathcal{L}} P$.

Hence if S is separable the set of $\omega$ such that $T_n(X^{\omega}_{n,1},\ldots,X^{\omega}_{n,n}) \rightarrow_w T_0$ is measurable. The former definition is equivalent to $T_n(X^{\omega}_{n,1},\ldots,X^{\omega}_{n,n}) \xrightarrow{\mathcal{L}} T_0$ a.s.

Analogously we talk about bootstrap in probability if

(1.2) $D^*(\mathcal{L}\ (T_n(X^{\omega}_{n,1},\ldots,X^{\omega}_{n,n}),\mathcal{L}\ (T_0)) \rightarrow 0$ in probability.

As before this definition is independent of the distance. Obviously the bootstrap a.u. implies the bootstrap in probability. As before we usualy indicate the bootstrap in probability as $T_n(X^{\omega}_{n,1},\ldots,X^{\omega}_{n,n}) \xrightarrow{\mathcal{L}} T_0$ in pr.

The above definitions imply, as usual, that $T_n(X^{\omega}_{n,1},\ldots,X^{\omega}_{n,n}) \xrightarrow{\mathcal{L}} c$ in pr. if and only if for each sequence $n'$ there is a further sequence $n''$ such that $T_{n''}(X^{\omega}_{n'',1},\ldots,X^{\omega}_{n'',n''}) \xrightarrow{\mathcal{L}} c$ a.u.

Formally we have two probability spaces $(\Omega,\mathcal{A}\ ,P)$ and $(\Omega',\mathcal{A}'\ ,P')$ as well as a metric space $(S,\mathcal{S})$ with the Borel $\sigma$-algebra. We have also functions $X_n : \Omega \times \Omega' \rightarrow S$ that are measurable for the product $\sigma - algebra$. We define

$X_n \xrightarrow{\mathcal{L}} X$ in pr. if $\beta^*(\mathcal{L}\ (X_n),\mathcal{L}\ (X)) \xrightarrow{P} 0$

$X_n \xrightarrow{\mathcal{L}} X$ a.u.. if $\beta^*(\mathcal{L}\ (X_n),\mathcal{L}\ (X)) \xrightarrow{P} 0$ a.u.

where $\beta$ is any metric of the weak convergence for probabilities in S.

We also define

$X_n \xrightarrow{P} X$ in pr. if $\alpha^*(\mathcal{L}\ (X_n),\mathcal{L}\ (X)) \xrightarrow{P} 0$

$X_n \xrightarrow{P} X$ a.u.. if $\alpha^*(\mathcal{L}\ (X_n),\mathcal{L}\ (X)) \xrightarrow{P} 0$ a.u.

where $\alpha$ is any metric of the convergence in probability for r.v. from $\Omega'$ to S.

All usual properties about the convergence carry over to this scheme. The next proposition collects some properties that we need later.

**Proposition 1.1** As before S is a metric space

(a) Let $X_n : \Omega \times \Omega' \to S$ be a random variable for the product $\sigma - algebra$. Then the following are equivalent:

$X_n \xrightarrow{P'} X$ in P-pr.

$X_n \xrightarrow{P \times P'} X$ and

for each $\epsilon > 0$ $EP'\{|X_n| > \epsilon\} \to 0$

(b) Let $X_n, Y_n : \Omega \times \Omega' \to S$ be r.v. for product $\sigma - algebra$. If $X_n \xrightarrow{\mathcal{L}} X$ in pr. (a.u.) and $Y_n \xrightarrow{\mathcal{L}} a$ in pr. (a.u.) where $a \in S$ then $X_n + Y_n \xrightarrow{\mathcal{L}} X + a$ in pr. (a.u.).

(c) Moreover if S is a linear metric space, $X_n : \Omega \times \Omega' \to S$ and $A_n : \Omega \times \Omega' \to \mathbf{R}$ are r.v. for the product $\sigma - algebra$, if $X_n \xrightarrow{\mathcal{L}} X$ in pr. (a.u.) and $A_n \xrightarrow{\mathcal{L}} a$ in pr. (a.u.) then $A_n X_n \xrightarrow{\mathcal{L}} aX$ in pr. (a.u.).

**Proof.** Since we can choose the distance we want for the convergence in probability for random variables $Y, Z : \Omega' \to \mathbf{R}$ we choose $\alpha(Y, Z) = \mathbf{E}|Y - Z| \wedge 1$ and for the probability measures on S $\beta(\mathbf{P}, \mathbf{Q}) = \sup \{|\mathbf{P}h - \mathbf{Q}h| : \text{where } h \in BL_1(S)\}$.

(a) From $(\mathbf{E}'|X_n - X| \wedge 1) \wedge 1 = \mathbf{E}'|X_n - X| \wedge 1$ we get that $X_n \xrightarrow{P'} X$ in P-pr. is equivalent to $X_n \xrightarrow{P \times P'} X$.

Note also that $\epsilon EP'\{|X_n - X| > \epsilon\} \leq \mathbf{E}\mathbf{E}'|X_n - X| \leq \epsilon + EP'\{|X_n - X| > \epsilon\}$. So that $X_n \xrightarrow{P \times P'} X$ is equivalent that for each $\epsilon > 0$ $EP'\{|X_n| > \epsilon\} \to 0$

(b) Note that $\beta(X_n + Y_n, X + Y) \leq \beta(X_n, X) + \beta(Y_n, Y)$.

(c) It is routine. □

# CHAPTER II

# THE BOOSTRAP OF THE MEAN WITH ARBITRARY BOOTSTRAP SAMPLE SIZE

## 1. Introduction.

The initial result about the bootstrap of the mean is Theorem 2.2 in Bickel-Freedman (1981). In our notation they showed that for random variables with values in $\mathbf{R}^d$ and finite second moment:

(1.1) $n^{-1/2}\sum_{i=1}^{n}(X_{ni}^{\omega}-\bar{X}_n(\omega))\xrightarrow{\mathcal{L}}\mathrm{N}(0,A)$ $\omega$–a.s.

where $A$ is the variance-covariance matrix of X and $\bar{X}_n(\omega)=n^{-1}\sum_{i=1}^{n}X_i$. It was believed that the bootstrap fails if there is no second moment. Babu (1984) showed that if $X_1,\ldots,X_n$ are symmetric stable random variables with index p with $1<p<2$ then, in spite of the fact that $n^{-1/p}\sum_{i=1}^{n}X_i\xrightarrow{\mathcal{L}}X$ it is not true that

(1.2) $n^{-1/p}\sum_{i=1}^{n}(X_{ni}^{\omega}-\bar{X}_n(\omega))\xrightarrow{\mathcal{L}}X$ $\omega$–a.s.

Another work about the bootstrap of the mean is in Athreya (1984, 1985 and 1987). He studied the bootstrap for i.i.d. r.v. taken from a stable law of order $\alpha$, $0<\alpha<2$. He determined how (1.2) is not true. He showed that for each x $\hat{\mathbf{P}}\{\sum_{i=1}^{n}(X_{ni}^{\omega}-\bar{X}_n(\omega))/n^{1/p}\le x\}\xrightarrow{\mathcal{L}}H(x)$ where H(x) is the distribution of a random measure. Besides that he showed that $\hat{\mathbf{P}}\{n^{-1/p}\sum_{i=1}^{n}(X_{ni}^{\omega}-\bar{X}_n(\omega))\le x\}$ converges to H(x) as stochastic processes in Skorohod space $D(-\infty,\infty)$. This result does not allow to recover (an approximation to) the law of $\bar{X}_n$.

Giné-Zinn (1989a) showed that if the bootstrap of the mean conveges $\omega$–a.s. then $\mathbf{E}X^2<\infty$ and that if the bootstrap of the mean converges in probability then X is in the domain of atraction of a normal random variable. Explicitly they showed:

**Theorem 1.1.** (Giné-Zinn 1989a) If there exist random variables $c_n(\omega), n \in \mathbf{N}$, a strictly increasing sequence $a_n \rightarrow \infty$ and a randon probability measure $\mu(\omega)$ nondegerate with positive probability such that $a_n^{-1}\sum_{i=1}^n X^\omega_{ni} - c_n(\omega) \xrightarrow{\mathcal{L}} \mu(\omega)$ $\omega$—a.s. then $a_n \cong n^{-1/2}$, $\mathrm{E}X^2 < \infty$ and $n^{-1/2}\sum_{i=1}^n (X^\omega_{ni} - \bar{X}_n(\omega)) \xrightarrow{\mathcal{L}} \mathrm{N}(0, Var\, X)$ a.s.

**Theorem 1.2.** (Giné-Zinn (1989a)) If there exist random variable $c_n(\omega)$, $n \in \mathbf{N}$, a strictly increasing sequence $a_n \rightarrow \infty$ and a random measure $\mu(\omega)$ nondegenerate with positive probability such that $a_n^{-1}\sum_{i=1}^n X^\omega_{ni} - c_n(\omega) \xrightarrow{\mathcal{L}} \mu(\omega)$ in pr. then $a_n = n^{1/2}L(n)$ where $L(n)$ is slowly varying and there exists $\sigma \neq 0$ such that $a_n^{-1}\sum_{i=1}^n (X_i - \mathrm{E}X) \xrightarrow{\mathcal{L}} \mathrm{N}(0, \sigma^2)$ and $a_n^{-1}\sum_{i=1}^n (X^\omega_{ni} - \bar{X}_n(\omega)) \xrightarrow{\mathcal{L}} N(0, \sigma^2)$ in pr.

These results show that the plain bootstrap cannot be implemented for stable limits. Several alternatives have been given to this problem. Athreya proposed to change the size of the bootstrap sample. He showed that it is possible to bootstrap the CLT of the mean in the stable case in probability when $m_n \rightarrow \infty$ and $m_n/n \rightarrow 0$. See Cor 3.6.

The facts that the techniques used in Giné-Zinn (1989a) are more adapted to the bootstrap of the mean and that Athreya's work did not solve completely the bootstrap in the case of a bootstrap size $m_n$ , induced us, E. Giné and I, to study the bootstrap of the mean in the $m_n$ case. This chapter is a revised exposition of our work ( Arcones-Giné (1989a) ). Our main results are:

(1) If X is in the domain of atraction of the normal law then it satisfies the bootstrap CLT in probability for al $m_n \rightarrow \infty$ (Theorem 3.2).

(2) If $\{m_n\}$ is a regular sequence such that $m_n(\log\log\ n)/n \rightarrow 0$, and X is in the domain of

atraction of a stable law then the bootstrap CLT holds a.s.; but it does not hold a.s. if $EX^2 = \infty$ and $\inf m_n(\log\log n)/n > 0$ (Section 4).

## 2. Heuristics and fundamental techniques.

First I am going to expose the way Giné-Zinn (1989a) tackle the problem of the bootstrap of the mean in the simplest situation. Suppose that we have a sample $X_1, \ldots, X_n$ from a random variable with finite second moment.

The classical CLT says $n^{-1/2}\sum_{i=1}^{n}(X_i - EX) \xrightarrow{\mathcal{L}} N(0, Var(X))$.

The bootstrap CLT says

(2.1) $m_n^{-1/2}\sum_{i=1}^{m_n}(X_{n\,i}^{\omega} - \bar{X}_n) \xrightarrow{\mathcal{L}} \mathrm{N}(0, Var(X))$ $\omega$-a.s. for any sequence $m_n \to \infty$.

Their way to prove it is to appeal to the CLT for triangular arrays conditionally in $\omega$. Given that the limit is normal we may use Corollary 4.8 b in page 63 Araujo-Giné (1980). We must check

(2.2) $m_n^{-1/2}(X_{n\,i} - \bar{X}_n)$ are centered

(2.3) $m_n^{-1}\sum_{i=1}^{m_n} E'(X_{n\,i} - \bar{X}_n)^2 \to Var(X)$ a.s. and

(2.4) for each $\epsilon > 0$ $m_n^{-1}\sum_{i=1}^{m_n} E'(X_{n\,i} - \bar{X}_n)^2 I_{|X_{n\,i} - \bar{X}_n| > \epsilon m_n^{1/2}} \to 0$ a.s.

(2.2) is obvious.

As far as (2.3) $m_n^{-1}\sum_{i=1}^{m_n} E'(X_{n\,i} - \bar{X}_n)^2 = n^{-1}\sum_{i=1}^{n} X_i^2 - (n^{-1}\sum_{i=1}^{n} X_i)^2 \to Var(X)$ a.s. because of the law of large numbers.

Finally since $\bar{X}_n \to EX$ a.s., for any $c > 0$ and n large

$m_n^{-1}\sum_{i=1}^{m_n}(X_i - \bar{X}_n)^2 I_{|X_i - \bar{X}_n| > 2\epsilon m_n^{1/2}} \leq n^{-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 I_{|X_i| \geq c} \to$
$EX^2 I_{|X| \geq c} - 2(EX)EXI_{|X| \geq c} + (EX)^2 P\{|X| \geq c\}$ a.s.

This expression can be made arbitrarily small for c large . So, (2.4) holds.

Bickel and Freedman not only got weak convergence a.s. but also convergence in the Mallows distance $d_2$. Next I will sketch some facts about the Mallows distance taken from their appendix. Let B be a Banach space with norm $||\cdot||$. Let $\Gamma_p$ be the set of probabilities $\mu$ on the Borel $\sigma$-field of B such that $\int ||x||^p d\mu(x) < \infty$. Then $d_p(\mu, \nu) = \inf\{||X - Y||_p$: over all the pairs of B-valued random variables with $\mathcal{L}\ (X) = \mu$ and $\mathcal{L}\ (Y) = \nu\}$. The importance of the Mallows distance rests on the equivalence of the following propositions, for $\alpha_n, \alpha \in \Gamma_p$

a) $d_p(\alpha_n, \alpha) \to 0$.

b) $\alpha_n \to \alpha$ and $\int ||x||^p d\alpha_n(x) \to \int ||x||^p d\alpha(x)$.

c) $\alpha_n \to \alpha$ and $||x||^p$ in uniformly $\alpha_n$-integrable.

d) $\int \phi\, d\alpha_n(x) \to \int \phi\, d\alpha(x)$ for every continuous $\phi$ such that $\phi(x) = O(||x||^p)$ at infinity.

In our case $n^{-1/2} \sum_{i=1}^n (X_i - EX) \xrightarrow{d_2} N(0, Var(X))$ because we have weak convergence by the CLT and $E(n^{-1/2} \sum_{i=1}^n (X_i - EX))^2 = Var(X)$. So for any continuous function $f : \mathbf{R} \to \mathbf{R}$ satisfying $f(x) = O(|x|^2)$ we have that $Ef(n^{-1/2} \sum_{i=1}^n (X_i - EX)) \to Ef(N(0, Var(X))$.

Bickel and Freedman (1981) showed that

$m_n^{-1/2} \sum_{i=1}^{m_n} (X_{ni}^\omega - \bar{X}_n) \xrightarrow{d_2} \mathrm{N}(0, Var(X))$ a.s.

The above proof gives it authomatically: it is just (2.3).

Another fundamental tool is the following lemma. See Giné-Zinn (1989a) or Chen and Rubin (1984).

**Lemma 2.1.** If $\mathrm{E}|X|^p = \infty$ and $0 < p' < p$, then

$(n^{-1} \sum_{i=1}^n |X_i|^{p'})^{1/p'} / (n^{-1} \sum_{i=1}^n |X_i|^p)^{1/p} \to 0$ a.s.

Moreover if $a_n \to \infty$ then

$(n^{-1} \sum_{i=1}^n |X_i|^{p'} I_{\{|X_i| \le a_n\}})^{1/p'} / (n^{-1} \sum_{i=1}^n |X_i|^p I_{\{|X_i| \le a_n\}})^{1/p} \to 0$ a.s.

## 3. The bootstrap CLT in probability of the mean.

The following theorem gives necessary conditions for the bootstrap in probability. Its proof closely follows Giné-Zinn (1989a). We use the notation $c_\tau \mathit{Pois}\,\pi$ for a generalized Poisson measure with Lévy measure $\pi$, as in Araujo-Giné (1980). If a sequence $\{c_n\}$ is nondecreasing and $c_n \to \infty$, we write $c_n \nearrow \infty$.

**Theorem 3.1.** Let X be a random variable for which there exist a sequence of positive integers $m_n \to \infty$, a sequence of positive real numbers $a_n \to \infty$, random variables $c_n(\omega)$ , n∈N, and a random probability measure $\mu(\omega)$, nondegenerate with positive probability, such that

(3.1) $a_n^{-1} \sum_{j=1}^{m_n} X_{nj}^{\omega} - c_n(\omega) \xrightarrow{\mathcal{L}} \mu(\omega)$ in probability.

Then:

(a) There are a Lévy measure $\pi$ and a real number $\sigma^2 \geq 0$ such that for all $\tau$ with $\pi\{-\tau,\tau\} = 0$

(3.1)' $\sum_{j=1}^{m_n} ( X_{nj}^{\omega} - n^{-1} \sum_{i=1}^{n} X_i(\omega) I_{|X_i(\omega)| \leq \tau a_n})/a_n \xrightarrow{\mathcal{L}} \mathrm{N}(0,\sigma^2) * c_\tau \mathit{Pois}\,\pi$ in probability.

(b) If c is a positive number and

(3.2) $b_{n,c} = a_n \qquad r_{n,c} = m_n \qquad \text{if} \quad m_n \leq cn$

(3.2)' $b_{n,c} = a_n(n/m_n)^{1/2} \qquad r_{n,c} = n \qquad \text{if} \quad m_n > cn$

then, for all $\tau$ with $\pi\{-\tau,\tau\} = 0$

(3.3) $\sum_{i=1}^{r_{n,c}} (X_i - \mathrm{E}X I_{|X| \leq \tau b_{n,c}} \xrightarrow{\mathcal{L}} \mathrm{N}(0,\sigma^2) * c_\tau \mathit{Pois}\,\pi$

(c) If $\limsup_{n\to\infty} m_n/n > 0$ then $\pi = 0$ and $\sigma^2 \neq 0$ in (a) and (b).

(d) If $\liminf_{n\to\infty} m_n/n > 0$ then X is in the domain of attraction of the normal law with norming constants $b_n = a_n(n/m_n)^{1/2}$, that is

(3.4) $b_n^{-1} \sum_{i=1}^{n} (X_i - \mathrm{E}X) \xrightarrow{\mathcal{L}} \mathrm{N}(0,\sigma^2)$.

(e) If (3.1) holds only along a subsequence $\{n'\} \subset \mathrm{N}$ with $\{m_{n'}\} \subset \{n'\}$, then (a)-(d) hold but with n' instead of n ( in (d) (3.4) holds for $n'$ but X is not neccessarily in the domain of attraction of $\mathrm{N}(0,\sigma^2)$ ).

**Proof.** If $X^2 < \infty$ this theorem is a consequence of Theorem 2.1 in Bickel-Freedman (1981). So, we assume $\mathbf{E}X^2 = \infty$. The proof that $\{a_n^{-1}X_{nj}^{\omega}\}$ is an infinitesimal system is similar to that of (2.3), a fact that will be used throughout without further mention. (3.1) holds if and only if every subsequence has a further subsequence along which the limit in (3.1) holds a.s.. Hence, $\mu(\omega)$ is a.s. infinitely divisible. Let $n'$ be such a subsequence. The fact that $a_n \to \infty$ readily gives, by the corresponding argument in Giné-Zinn (1989a), that the Lévy measure $\pi(\omega)$ of $\mu(\omega)$ is a.s. a fixed Lévy measure $\pi$. Moreover, if D is a countable set of points $\delta$ dense in $\mathbf{R}^+$ such that $\pi\{-\delta,\delta\} = 0$ then the following limits hold for all $\delta \in D$ almost surely by the general CLT in **R**

(3.5) $(m_{n'}/n') \sum_{i=1}^{n'} I_{X_i > \delta a_{n'}} \to \pi(\delta,\infty)$ $\omega$-a.s.

(3.5)' $(m_{n'}/n' \sum_{i=1}^{n'} I_{X_i < -\delta a_{n'}} \to \pi(-\infty,-\delta)$ $\omega$-a.s. and

(3.6) $(m_{n'}/n'a_{n'}^2) \sum_{i=1}^{n'} X_i^2 I_{|X_i| \le \delta a_{n'}} - (m_{n'}/n'^2 a_{n'}^2)(\sum_{i=1}^{n'} X_i I_{|X_i| \le \delta a_{n'}})^2 \to$

$\sigma^2(\omega) + \int_{-\delta}^{\delta} x^2 \, d\pi(x)$ $\omega$-a.s.

where $\sigma^2(\omega)$ is the variance of the normal component of $\mu(\omega)$. Since $\mathbf{E}X^2 = \infty$ by Lemma 2.1,

(3.7) $(n^{-1} \sum_{i=1}^{n} X_i I_{|X_i| \le \delta_n})^2 \,/\, n^{-1} \sum_{i=1}^{n} X_i^2 I_{|X_i| \le \delta_n} \to 0$ a.s.

Then (3.6) becomes

(3.8) $(m_{n'}/n'a_{n'}^2) \sum_{i=1}^{n'} X_i^2 I_{|X_i| \le \delta a_{n'}} \to \sigma^2(\omega) + \int_{-\delta}^{\delta} x^2 \, d\pi(x)$ a.s .

In particular, $m_{n'}/n'a_{n'}^2 \to 0$ (even $m_{n'}/a_{n'}^2 \to 0$ by the law of the large numbers) and therefore $\sigma^2(\omega)$ is a tail random variable, hence a.s. constant, say $\sigma^2$. We have thus shown that $\sigma(\omega)$ is a tail random variable, hence a.s. constant, say $\sigma^2$. We have thus shown that $\mu(\omega)$ is a.s. a (possibly random) shift of $\mu = \mathrm{N}(0,\sigma^2) * c_\tau Pois\, \pi$. Moreover (3.5) and (3.6) show, by the general CLT in **R**, that (3.1)' holds a.s along n'. Therefore (3.1)' holds.

We will prove (b) for $c = 1$, the case of general $c$ being entirely similar. Set $b_n = b_{n,1}$, and $r_n = r_{n,1} = m_n \wedge n$. Consider a subsequence for which (3.1) converges a.s. Then either $m_n/n \leq 1$ infinitely often along this sequence, or $m_n/n > 1$ i.o. or both. So, we may specialize to two types of subsequences $n'$ for which (2.1) holds a.s.:those satisfying $m_{n'}/n' \leq 1$, and those for which $m_{n'}/n' > 1$. In the first case the summands in (3.5) and (3.6) are bounded and therefore their expected values converge to the expected values of their respective limits (by e.g.Acosta-Giné (1979), theorem 3.2 or exercises 13 and 14, pages 69, 70 in Araujo-Giné (1980) ). Thus, we obtain for all $\delta \in \mathbf{D}$, $m_{n'}\mathbf{P}\{X > \delta a_{n'}\} \to \pi(\delta, \infty)$, $m_{n'}\mathbf{P}\{X < -\delta a_{n'}\} \to \pi(-\infty, -\delta)$ and $(m_{n'}/a_{n'}^2)EX^2 I_{|X| \leq \delta a_{n'}} \to \sigma^2 + \int_{-\delta}^{\delta} x^2 \, d\pi(x)$. These three limits imply [recall $EX^2 = \infty$ and (3.7)] by the general CLT in $\mathbf{R}$ that the the limit (3.3) holds along $\{n'\}$ ($b_{n'} = a_{n'}$ in this case). Suppose now $m_{n'}/n' > 1$ and that (3.1) holds a.s. along $n'$. Then, by taking a further subsequence if necessary, we may assume $m_{n'}/n' \to c \in [1, \infty]$. If $c = \infty$ then (3.5) implies $\sum_{i=1}^{n'} I_{|X_i| > \delta a_{n'}} \to 0$ a.s. If $c < \infty$ then the corresponding argument ( on binomial limits) in the proof of Theorem 1 in Giné-Zinn (1989a) gives also that $\sum_{i=1}^{n'} I_{|X_i| > \delta a_{n'}} \to 0$ a.s. Then $\sum_{i=1}^{n'} I_{|X_i| > \delta a_{n'}} = 0$ eventually a.s. (since this sequence is integer valued and tends to 0), which implies that the limits in (3.5) are 0 a.s. that is, $\pi = 0$. This argument already proves (c). Then, since $\sum_{i=1}^{n'} X_i^2 I_{|X_i| > \delta a_{n'}} = 0$ eventually a.s., (3.8) becomes

(3.9) $\quad (m_{n'}/n' a_{n'}^2) \sum_{i=1}^{n'} X_i^2 \to \sigma^2$ a.s.

We can apply the converse CLT to (3.9) and obtain [recall $b_{n'} = a_{n'}(n'/m_{n'})^{1/2}$ in this case] that for all $\delta > 0$,

(3.10) $\quad n'\mathbf{P}\{|X| > \delta b_{n'}\} \to 0 \qquad (n'/b_{n'}^2)\mathbf{E}X^2 I_{|X| \leq \delta b_{n'}} \to \sigma^2.$

But (3.10) implies by the CLT in $\mathbf{R}$ (recall $\mathbf{E}X^2 = \infty$),

(3.11) $\quad b_{n'}^{-1} \sum_{i=1}^{n'} (X_i - \mathbf{E}X I_{|X| \leq \delta b_{n'}}) \xrightarrow{\mathcal{L}} \mathbf{N}(0, \sigma^2).$

This argument already proves (d) since it gives (3.4) along a subsequence of every subsequence.

We have thus proved that every subsequence has a further subsequence along which the limit (3.3) holds. Hence, (3.3) holds. (a)-(d) are proved. The proof of (e) follows exactly along the same lines since nothing in the above arguments depends on the sequence $\{n\}$ being all of $\mathbf{N}$. □

Now we consider the case of X in the domain of atraction of a normal law. There are two reasons for this: one is its importance, and the other is that in this case we have regular variation as an additional convenient tool. Theorem 3.2 was obtained by Athreya (1987) in the special case $m_n = n$ and by Csörgõ and Mason (1989) for $0 < c_1 \le m_n/n \le c_2 < \infty$. We extend their results to arbitrary $\{m_n\}$ and study some related topics.

Concretely we assume that there are constants $b_n$ so that

(3.12) $\sum_{i=1}^{n}(X_i - \mathrm{E}X)/b_n \xrightarrow{\mathcal{L}} \mathrm{N}(0,1)$.

Now I enumerate some known facts about $X$. A reference for all this facts is section 6 in chapter II of Araujo-Giné (1980). Define $U(x) = EX^2 I_{|X|\le x}$. Then

(3.13) $U(\cdot)$ is slowly varying.

(3.14) $nb_n^{-2}U(\tau b_n) \to 1$ for any $\tau > 0$.

Note that for each Borel measurable function satistying $f(0) = 0$

(3.15) $Ef(|X|) = \int_0^\infty x^{-2} f(x)\, dU(x)$ in the sense that if one of parts of the equality is defined the other is defined and they are equal.

Using (3.15), integration by parts and lemma II.6.15 in Araujo-Giné (1980) we get that

(3.16) $x^{2-p}E|X|^p I_{|X|\ge x}/U(x) \to 0$ for $0 < p < 2$

(3.17) $x^{2-p}E|X|^p I_{|X|\le x}/U(x) \to 0$ for $2 < p$.

Then for (3.14) and (3.16)-(3.17) we get

(3.18) $nb_n^{-p}E|X|^p I_{|X|\ge \tau b_n} \to 0$ for $p < 2$ and any $\tau > 0$

(3.19) $nb_n^{-p}E|X|^p I_{|X|\le \tau b_n} \to 0$ for $p > 2$ and any $\tau > 0$

**Theorem 3.2.** Let $X$ be as in (3.12). Let $m_n \to \infty$. For $c \in (0, \infty)$ fixed, let

$a_n = b_{m_n}$ if $m_n \leq cn$ and

$a_n = b_n(m_n/n)^{1/2}$ if $m_n > cn$ then

(3.20) $\sum_{i=1}^{m_n}(X^{\omega}_{ni} - \bar{X}_n(\omega))/a_n \xrightarrow{\mathcal{L}} \mathrm{N}(0,1)$ in probability.

**Proof.** The case $\mathbf{E}X^2 < \infty$ was done in the section 2. So, we assume $\mathbf{E}X^2 = \infty$. Using the central limit theorem for triangular arrays conditionally on $\omega$ we must show

(3.21) $(m_n/a_n^2)[n^{-1}\sum_{i=1}^n X_i^2 I_{|X_i| \leq \tau a_n} - (n^{-1}\sum_{i=1}^n X_i I_{|X_i| \leq \tau a_n})^2] \xrightarrow{P} 1$ for all $\tau > 0$

(3.22) $(m_n/n)\sum_{i=1}^n I_{|X_i| \geq \tau a_n} \xrightarrow{P} 0$ for all $\tau > 0$ and

(3.23) $(m_n/na_n)\sum_{i=1}^n X_i I_{|X_i| \geq \tau a_n} \xrightarrow{P} 0$ for all $\tau > 0$ .

Note that for lemma 2.1 (3.21) is equivalent to

(3.24) $(m_n/na_n^2)\sum_{i=1}^n X_i^2 I_{|X_i| \leq \tau a_n} \xrightarrow{P} 1$ for all $\tau > 0$ .

First we consider the case $m_n \leq cn$

$E(m_n/na_n^2)\sum_{i=1}^n X_i^2 I_{|X_i| \leq \tau a_n} = (m_n/b_{m_n}^2)U(\tau b_{m_n}) \to 1$ for (3.14).

$\mathrm{Var}((m_n/na_n^2)\sum_{i=1}^n X_i^2 I_{|X_i| \leq \tau a_n}) \leq (m_n^2/nb_{m_n}^4)E|X|^4 I_{|X| \leq \tau b_{m_n}} \to 0$ for (3.18).

$E|(m_n/n)\sum_{i=1}^n I_{|X_i| \geq \tau a_n}| \leq m_n P\{|X| \geq \tau b_{m_n}\} \to 0$ for (3.19).

$E(m_n/na_n)\sum_{i=1}^n X_i I_{|X_i| \geq \tau a_n}| \leq (m_n/b_{m_n})E|X| I_{|X| \geq \tau b_{m_n}} \to 0$ for (3.18).

If $m_n \geq cn$ $E\sum_{i=1}^n I_{|X_i| \geq \tau a_n} \leq nP\{|X| \geq c\tau b_n\} \to 0$. So $P\{\sum_{i=1}^n I_{|X_i| \geq \tau a_n} = 0\} \to 1$. This implies (3.22), (3.23) and that $P\{\sum_{i=1}^n X_i^2 I_{|X_i| \leq \tau a_n} \neq \sum_{i=1}^n X_i^2\} \to 0$. Hence to show (3.24) we need to show

(3.25) $b_n^{-2}\sum_{i=1}^n X_i^2 \xrightarrow{P} 1$

But this follows directly from the CLT using (3.13)-(3.19). □

Now the question that we pose is what can we say about

(3.26) $a_n^{-1}\sum_{i=1}^{m_n}(X^{\omega}_{n\,i} - \bar{X}_n(\omega)) \xrightarrow{d_p} \mathrm{N}(0,1)$ in probability.

Note that $Var'(\sum_{i=1}^{m_n}(X^{\omega}_{n\,i} - \bar{X}_n(\omega))/a_n) = (m_n/a_n^2)(n^{-1}\sum_{i=1}^{n} X_i^2 - (n^{-1}\sum_{i=1}^{n} X_i)^2)$. By Lemma 2.1 this expression is convergent equivalent to $(m_n/na_n^2)\sum_{i=1}^{n} X_i^2$. Since $b_n^{-2}\sum_{i=1}^{n} X_i^2 \xrightarrow{P} 1$, we must consider the convergence of $m_n b_n^2/na_n^2$. In the case of $m_n \geq cn$ this is one. In the case $m_n \leq cn$ $m_n b_n^2/na_n^2 = m_n b_n^2/nb_{m_n}^2$. Note that $U(b_n)/b_n^2 \cong 1/n$. So $b_n^2/n \to \infty$. Hence there is a sequence $m_n \to \infty$ such that $m_n \leq n$ and $m_n b_n^2/nb_{m_n}^2 \to \infty$. Next I will show that for any $0 < p < 2$ (3.26) holds.

**Theorem 3.3.** Under the conditions of theorem 3.2 we have

(a) $a_n^{-1}\sum_{i=1}^{m_n}(X^{\omega}_{n\,i} - \bar{X}_n(\omega)) \xrightarrow{d_p} \mathrm{N}(0,1)$ in pr. for any $0 < p < 2$.

(b) Moreover if $\liminf m_n/n > 0$ then for any $t \in \mathbf{R}$

$E'\exp(t\, a_n^{-1}\sum_{i=1}^{m_n}(X^{\omega}_{n\,i} - \bar{X}_n)) \to E\exp(tg)$ in pr. where g is a $\mathrm{N}(0,1)$ v.r.. In particular for any $0 < p < \infty$ $a_n^{-1}\sum_{i=1}^{m_n}(X^{\omega}_{n\,i} - \bar{X}_n(\omega)) \xrightarrow{d_p} \mathrm{N}(0,1)$ in pr.

**Proof of a.** From section 2, I may restrict to the case $EX^2 = \infty$ and for the last paragraph, I may restrict to $m_n/n \leq c$. Hence I must show for $EX^2 = \infty$ and $m_n/n \leq c$ that (3.26) holds for $0 < p < 2$. First I prove that $E'|a_n^{-1}\sum_{i=1}^{m_n}(X^{\omega}_{n,i} - \bar{X}_n)|^p$ is bounded in probability for $1 < p < 2$.

Call $Y_{n,i}$ to an independent i.i.d. copy of $X_{n,i}$ then

(3.27) $E'|a_n^{-1}\sum_{i=1}^{m_n}(X^{\omega}_{n,i} - \bar{X}_n)|^p \leq E'|a_n^{-1}\sum_{i=1}^{m_n}(X^{\omega}_{n,i} - Y^{\omega}_{n,i})|^p =$

$E'|a_n^{-1}\sum_{i=1}^{m_n}\epsilon_i(X^{\omega}_{n,i} - Y^{\omega}_{n,i})|^p \leq 2^p E'|a_n^{-1}\sum_{i=1}^{m_n}\epsilon_i X^{\omega}_{n,i}|^p \leq 2^p p^{p/2} E'|a_n^{-2}\sum_{i=1}^{m_n}(X^{\omega}_{n,i})^2|^{p/2} \leq$

$2^p p^{p/2} E'|a_n^{-2}\sum_{i=1}^{m_n} X^2_{n,i} I_{|X_{n,i}| \leq c_n}|^{p/2} + 2^p p^{p/2} E'|a_n^{-2}\sum_{i=1}^{m_n} X^2_{n,i} I_{|X_{n,i}| \geq c_n}|^{p/2} \leq$

$2^p p^{p/2}[E'|a_n^{-2}\sum_{i=1}^{m_n}(X^{\omega}_{n,i})^2 I_{|X_{n,i}| \leq c_n}|]^{p/2} + 2^p p^{p/2} E'|a_n^{-2}\sum_{i=1}^{m_n} X^2_{n,i} I_{|X_{n,i}| \geq c_n}|^{p/2} \leq$

$2^p p^{p/2}[(m_n/a_n^2)E'|(X^2_{n,i} I_{|X_{n,i}| \leq c_n}|]^{p/2} + 2^p p^{p/2}(m_n/a_n^p)E'|X_{n,i}|^p I_{|X_{n,i}| \geq c_n}$

$2^p p^{p/2}[(m_n/na_n^2)\sum_{i=1}^n |X_i|^2 I_{|X_i|\le c_n}]^{p/2} + 2^p p^{p/2}(m_n/na_n^p)\sum_{i=1}^n |X_i|^p I_{|X_i|\ge c_n}$.

Note that this inequality holds for any bootstrap sample. Take $c_n = b_{m_n}$. We must show that $(m_n/nb_{m_n}^2)\sum_{i=1}^n |X_i|^2 I_{|X_i|\le b_{m_n}}$ and $(m_n/nb_{m_n}^p)\sum_{i=1}^n |X_i|^p I_{|X_i|\ge b_{m_n}}$ are bounded in probability. Note that $E(m_n/nb_{m_n}^2)\sum_{i=1}^n |X_i|^2 I_{|X_i|\le b_{m_n}} = (m_n/b_{m_n}^2)EX^2 I_{|X|\le b_{m_n}} \to 1$ and

$E(m_n/nb_{m_n}^p)\sum_{i=1}^n |X_i|^p I_{|X_i|\ge b_{m_n}} = (m_n/b_{m_n}^p)EX^p I_{|X|\ge b_{m_n}} \to 0$.

Given that for any $0 < p < p'$ and $c > 0$

$d_p(\mu_n, \mu) \le (1 + pc^{p-1})||\mu_n - \mu||_{BL_1} + (c^{(p-p')/p}\mu_n|x|^{p'})^{1/p'\vee 1} + (c^{(p-p')/p}\mu|x|^{p'})^{1/p'\vee 1}$. So

$a_n^{-1}\sum_{i=1}^{m_n}(X_{n\,i}^{\omega} - \bar{X}_n(\omega)) \xrightarrow{BL_1}$ N(0,1) in pr. and

lim $_{M\to\infty}$ lim $\sup_{n\to\infty} P\{E'|a_n^{-1}\sum_{i=1}^{m_n}(X_{n\,i}^{\omega} - \bar{X}_n(\omega))|^p > M\} = 0$ for all $0 < p < 2$ implies $a_n^{-1}\sum_{i=1}^{m_n}(X_{n\,i}^{\omega} - \bar{X}_n(\omega)) \xrightarrow{d_p}$ N(0,1) in pr. for all $0 < p < 2$. □

**Proof of b** By symmetrization

$E' \exp(t\, a_n^{-1}\sum_{i=1}^{m_n}(X_{n\,i}^{\omega} - \bar{X}_n(\omega))) \le E' \exp(2t\, a_n^{-1}\sum_{i=1}^{m_n} \epsilon_i X_{n\,i})$

By the subgaussian inequality

$E' \exp(2t\, a_n^{-1}\sum_{i=1}^{m_n} \epsilon_i X_{n\,i}) \le E' \exp(2t^2\, a_n^{-2}\sum_{i=1}^{m_n} X_{n\,i}^2) = [n^{-1}\sum_{i=1}^n \exp(2t^2 X_i^2 a_n^{-2})]^{m_n}$

Note that $\max_{i\le n}|X_i|a_n^{-2} \le c^{-2}\max_{i\le n}|X_i|b_n^{-2} \xrightarrow{P} 0$

So $n^{-1}\sum_{i=1}^n \exp(2t^2 X_i^2 a_n^{-2}) \xrightarrow{P} 1$ and

$m_n \log[n^{-1}\sum_{i=1}^n \exp(2t^2 X_i^2 a_n^{-2})] \cong (m_n/n)\sum_{i=1}^n[\exp(2t^2 a_n^{-2} X_i^2) - 1] \cong$

$(m_n/a_n^2 n)\sum_{i=1}^n 2t^2 X_i^2 \xrightarrow{P} 2t^2$. □

**Remark 3.4.** (Random scaling). In the notation of the theorem 3.2 if $\hat{a}_n^{\omega}(\omega')$ is a random variable so that $\hat{a}_n^{\omega}(\omega')/a_n \xrightarrow{P'} 1$ in pr, then

(3.28) $\sum_{i=1}^{m_n}(X_{n\,i}^{\omega} - \bar{X}_n(\omega))/\hat{a}_n^{\omega}(\omega') \to N(0,1)$ in pr.

A sequence of random variable satisfying this condition is

(3.29) $\hat{a}_n^\omega = [(m_n/n)\sum_{i=1}^n X_i^2(\omega)]^{1/2}$ if $m_n \geq n$ and

(3.29') $\hat{a}_n^\omega$ =average over all $\binom{n}{m_n}$ combinations $1 \leq j_1 < \cdots < j_{m_n} \leq n$ of

$[\sum_{i=1}^{m_n}(X_{j_i} - m_n^{-1}\sum_{i=1}^{m_n} X_{j_i})^2]^{1/2}$ if $m_n \leq n$.

If $m_n \geq n$ $[\hat{a}_n^\omega(\omega')/a_n]^2 = b_n^{-2}\sum_{i=1}^n X_i^2 - n\, b_n^{-2}(n^{-1}\sum_{i=1}^n X_i)^2 \xrightarrow{P} 1$ by the CLT.

If $m_n \leq n$

$P\{|\hat{a}_n^\omega(\omega')a_n^{-1} - 1| > \epsilon\} \leq \epsilon^{-1}E|\hat{a}_n^\omega(\omega')a_n^{-1} - 1| \leq \epsilon^{-1}E|[b_{m_n}^{-2}\sum_{i=1}^{m_n}(X_i - m_n^{-1}\sum_{i=1}^{m_n} X_i)^2]^{1/2} - 1|$.

We have that $b_n^{-2}\sum_{i=1}^n(X_i - n^{-1}\sum_{i=1}^n X_i)^2 \xrightarrow{P} 1$.

Using the argument in (3.27) for $0 < p < 1$ $E|b_n^{-2}\sum_{i=1}^n X_i^2|^p \leq (nb_n^{-2}E|X|^2 I_{|X| \leq b_n})^p + nb_n^{-2p}E|X|^2 I_{|X| \geq b_n} \rightarrow 1$. So $E|(b_n^{-2}\sum_{i=1}^n X_i^2)^{1/2} - 1| \rightarrow 1$.

We also can take the bootstrap choice

(3.30) $\hat{a}_n^\omega = (\sum_{i=1}^n(X_{n\,i}^\omega - \hat{X}_n)^2)^{1/2}$ if $m_n \leq n$.

(3.30') $\hat{a}_n^\omega$ = average over all $\binom{m_n}{n}$ possible combinations $1 \leq j_1 < \cdots j_n \leq m_n$ of

$[\sum_{i=1}^n(X_{n,j_i} - n^{-1}\sum_{i=1}^n X_{n,j_i})^2]^{1/2}$ if $m_n \geq n$.

If $m_n \leq n$ we must show $b_{m_n}^{-2}\sum_{i=1}^n(X_{n\,i}^\omega - \hat{X}_n)^2 \xrightarrow{P} 1$. This can be shown again by the CLT, conditinally on $\omega$. (It holds for any $m_n \rightarrow \infty$).

If $m_n \geq n$, by proposition 1 in the chapter I it is enough to show $P'\{|\hat{a}_n^\omega(\omega')/a_n - 1| > \epsilon\} \rightarrow 0$ in pr. As before $P'\{|\hat{a}_n^\omega(\omega')a_n^{-1} - 1| > \epsilon\} \leq \epsilon^{-1}E'|[b_{m_n}^{-2}\sum_{i=1}^{m_n}(X_{n,i} - \bar{X}_n)^2]^{1/2} - 1|$. This holds because $b_{m_n}^{-2}\sum_{i=1}^n(X_{n\,i}^\omega - \hat{X}_n)^2 \xrightarrow{P} 1$ in pr. and that for any $0 < p < 1$ $E'[b_{m_n}^{-2}\sum_{i=1}^{m_n}(X_{n,i} - \bar{X}_n)^2]^p$ is bounded in probability as we saw in the theorem 3.3.

Finally we consider the general case of X in the domain of partial attraction of an infinitely divisible law.

**Theorem 3.5.** Let X satisfy

(3.31) $\sum_{i=1}^{n'}(X_i - \mathbf{E}XI_{|X|\le\tau a_{n'}})/a_{n'} \xrightarrow{\mathcal{L}} \mathbf{N}(0,\sigma^2) * c_\tau\, Pois\,\pi$

for some $n' \to \infty$ , $a_{n'} \to \infty$, $\sigma^2 \geq 0$ and a Lévy measure $\pi$ (possibly 0), with $\tau$ such that $\pi\{-\tau,\tau\} = 0$. Let $m_{n'} \to \infty$ , $\{m_{n'}\} \subset \{n'\}$. Then if $m_{n'}/n' \to 0$ or if $\pi = 0$ and $\sup_{n'} m_{n'}/n' < \infty$,

(3.32) $\sum_{j=1}^{m_{n'}}( X^\omega_{n'j} - n'^{-1}\sum_{i=1}^{n'} X_i(\omega)I_{|X_i(\omega)\le\tau a_{m_{n'}}})/a_{m_{n'}} \xrightarrow{\mathcal{L}} \mathbf{N}(0,\sigma^2) * c_\tau\, Pois\,\pi$ in probability.

**Proof.** Here again, we can assume $EX^2 = \infty$. Let us first consider the case $m_{n'}/n' \to 0$. By the usual arguments ( with subsequences and the CLT ) it suffices to show that for every $\delta \in \mathbf{D}$,

(3.33) $m_{n'}\hat{\mathbf{P}}\{X^\omega_{n'1} > \delta a_{m_{n'}}\} = (m_{n'}/n)(\sum_{i=1}^{n'} I_{X_i>\delta a_{m_{n'}}}) \to \pi(\delta,\infty)$ in pr.

(3.34) $m_{n'}\hat{\mathbf{P}}\{X^\omega_{n'1} < -\delta a_{m_{n'}}\} = (m_{n'}/n')(\sum_{i=1}^{n'} I_{X_i<-\delta a_{m_{n'}}})$

(3.35) $(m_{n'}/a^2_{m_{n'}})\hat{V}ar(X^\omega_{n'1}I_{|X^\omega_{n',1}|\le\delta a_{m_{n'}}}) \to \sigma^2 + \int_{-\delta}^{\delta} x^2\, d\pi(x)$ in pr.

By lemma (3.35) is equivalent to

(3.36) $(m_{n'}/n'a^2_{m_{n'}})\sum_{i=1}^{n'} X_i^2 I_{|X_i|\le\delta a_{m_{n'}}}) \to \sigma^2 + \int_{-\delta}^{\delta} x^2\, d\pi(x)$ in pr.

(3.33) follows from the following two limits which hold by the converse CLT applied to (3.31) and $m_{n'}/n' \to 0$.

$\mathbf{E}[\,(m_{n'}/n')\ \sum_{i=1}^{n'} I_{X_i>\delta a_{n'}}\,] = m_{n'}\mathbf{P}\{X > \delta a_{m_{n'}}\} \to \pi(\delta,\infty)$ and

$Var[\,(m_{n'}/n')\ \sum_{i=1}^{n'} I_{X_i>\delta a_{n'}}] \leq (m^2_{n'}/n')\mathbf{P}\{X > \delta a_{m_{n'}}\} \to 0$

(3.34) is proved in exactly the same way. Now the proof of (3.36) is different in each case. Let us first consider the case $m_{n'}/n' \to 0$.

$\mathbf{E}[(m_{n'}/n'a^2_{m_{n'}})(\sum_{i=1}^{n'} X_i^2 I_{|X|\le\delta a_{m_{n'}}}\}] = (m_{n'}/a^2_{m_{n'}})\mathbf{U}(\delta a_{m_{n'}}) \to \sigma^2 + \int_{-\delta}^{\delta} x^2\, d\pi(x)$in pr. and

$Var[(m_{n'}/n'a^2_{m_{n'}})(\sum_{i=1}^{n'} X_i^2 I_{|X|\le\delta a_{m_{n'}}}] \leq (m^2_{n'}/n'a^4_{m_{n'}})\mathbf{E}X^4 I_{|X|\le\delta a_{m_{n'}}} \leq$

$\leq (m_{n'}/n')\,(\delta^2 m_{n'}/a^2_{m_{n'}})U(\delta a_{m_{n'}}) \to 0.$

Let us assume $\sup m_{n'}/n' \leq c < \infty$ and $\pi = 0$. Call

$Y_{n'}(\delta) = (m_{n'}/n'a^2_{m_{n'}}) \sum_{i=1}^{n'} X_i^2 I_{|X_i| \le \delta a_{m_{n'}}}$. Note that for all $\tau > 0$

(3.37) $\mathbf{E}Y_{n'}(\tau) = (m_n/a^2_{m_{n'}})U(\tau a_{m_{n'}}) \to \sigma^2$.

For $0 < \epsilon < \delta$

(3.38) $\mathbf{E}|Y_{n'}(\delta) - Y_{n'}(\epsilon)| = \mathbf{E}(m_{n'}/n'a^2_{m_{n'}} \sum_{i=1}^{n'} X_i^2 I_{\epsilon a_{m_{n'}} < |X| < \delta a_{m_{n'}}}$

$\le \delta^2 m_{n'} \mathbf{P}\{|X| > \epsilon a_{m_{n'}}\} \to 0$

Note also $Var((Y_{n'})(\epsilon)) \le (m^2_{n'}/n'a^4_{m_{n'}})\mathbf{E}X^4 I_{|X| < \epsilon a_{m_{n'}}} \le \epsilon^2 (m_{n'}/n')^2 (m_{n'}/a^2_{m_{n'}}) U(\epsilon a_{m_{n'}})$. So

(3.39) $\limsup Var(Y_{n'}(\epsilon)) \le \epsilon^2 c^2 \sigma^2$

Therefore, given $\mu > 0$, for $n$ large

(3.40) $\mathbf{P}(|Y_n(\delta) - \sigma^2| > \mu) \le \mathbf{P}(|Y_n(\delta) - Y_n(\epsilon)| > \mu/3) + \mathbf{P}(|Y_n(\epsilon) - \mathbf{E}Y_n(\epsilon)| > \mu/3) \le$

$(3/\mu)\mathbf{E}|Y_n(\delta) - Y_n(\epsilon)| + (9/\mu^2)Var(Y_n(\epsilon))$.

By (3.37)-(3.38) and (3.39) $\limsup \mathbf{P}(|Y_n(\delta) - \sigma^2| > \mu) \le 9\epsilon^2 c^2 \sigma^2/\mu^2$ for all $\epsilon > 0$. Hence (3.36) holds. □

Now I consider the case of stable limit. In the notation of Araujo-Giné (pages 49 and 80) $\theta$ has a non-normal stable law if and only if there are constants $\tau > 0$, $c_1 \ge 0$, $c_2 \ge 0$ and $0 < \alpha < 2$ such that $\mathcal{L}(\theta) = c_\tau$ Pois$\mu$ where

$d\mu(c_1, c_2, \alpha)(x) = c_1 x^{-1-\alpha} dx$ if $x > 0$

$d\mu(c_1, c_2, \alpha)(x) = c_2 |x|^{-1-\alpha} dx$ if $x < 0$.

If $X$ belongs to the domain of atraction of a stable law with exponent $\alpha$ $0 < \alpha < 2$, i.e. there are constants $c_n$ and $b_n$ so that

(3.41) $b_n^{-1} \sum_{i=1}^{n} (X_i - c_n) \to c_\tau$ Pois$\mu$ then all the followings facts are true

$nP\{X > b_n\} \to c_1/\alpha \qquad nP\{X < -b_n\} \to c_2/\alpha$

$P\{X > x\}$ and $P\{X < -x\}$ are regularly varying of order $-\alpha$.

$nb_n^{-2} U(b_n) \to (c_1 + c_2)/(2 - \alpha)$.

U(x) is regularly varying of order $2-\alpha$.

$x^{2-p}E|X|^p I_{|X|\geq x}/U(x) \to (2-\alpha)/(\alpha-p)$ for $p<\alpha$.

$x^{2-p}E|X|^p I_{|X|\leq x}/U(x) \to (2-\alpha)/(p-\alpha)$ for $p>\alpha$.

(3.42) $nb_n^{-p}E|X|^p I_{|X|\geq b_n} \to (c_1+c_2)/(\alpha-p)$ for $p<\alpha$.

(3.43) $nb_n^{-p}E|X|^p I_{|X|\leq b_n} \to (c_1+c_2)/(p-\alpha)$ for $p>\alpha$.

The previous theorem immediately gives:

**Corollary 3.7.** (Athreya (1987)) Let $\theta$ be a non-degenerate $\alpha$-stable random variable , $\alpha<2$, and let X be in its domain of attraction, with normig constants $b_n$ that is

(3.44) $b_n^{-1}\sum_{i=1}^{n}(X_i - \mathbf{E}XI(|X|\leq \tau b_n))/b_n\ ) \to_w \theta$

where $\tau$ can be zero if $\alpha<1$ and $+\infty$ if $\alpha>1$. Then if $m_n/n \to 0$,

(3.45) $b_{m_n}^{-1}\sum_{j=1}^{m_n}(\ X_{nj}^{\omega} - n^{-1}\sum_{i=1}^{n} X_i(\omega)I(|X_i(\omega)|\leq \tau b_{m_n})\ ) \xrightarrow{\mathcal{L}} \theta$ in probability.

**Proof.** We need to make some considerations in the centerings. Fix $0<\tau<\infty$ we must show that

$(m_n/nb_{m_n})\sum_{i=1}^{n}(X_iI_{|X_i|>\tau b_{m_n}} - EXI_{|X|>\tau b_{m_n}}) \xrightarrow{P} 0$ for $1<\alpha<2$ and

$(m_n/nb_{m_n})\sum_{i=1}^{n}(X_iI_{|X_i|\leq\tau b_{m_n}} - EXI_{|X|\leq\tau b_{m_n}}) \xrightarrow{P} 0$ for $0<\alpha<1$.

For $1<\alpha<2$ take $1<p<\alpha$. For symmetrization and Kintchine

$E|(m_n/nb_{m_n})\sum_{i=1}^{n}(X_iI_{|X_i|>\tau b_{m_n}} - EXI_{|X|>\tau b_{m_n}})|^p \leq E|(m_n^2/n^2b_{m_n}^2)\sum_{i=1}^{n}X_i^2I_{|X_i|>\tau b_{m_n}}|^{p/2} \leq$
$E(m_n^p/b_{m_n}^p n^p)\sum_{i=1}^{n}|X_i|^pI_{|X_i|>\tau b_{m_n}} = (m_n/n)^{p-1}(m_n/b_{m_n}^p)E|X_i|^pI_{|X_i|>\tau b_{m_n}} \to 0$

For $0<\alpha<1$

$E[(m_n/nb_{m_n})\sum_{i=1}^{n}(X_iI_{|X_i|\leq\tau b_{m_n}} - EXI_{|X|\leq\tau b_{m_n}})]^2 \leq (m_n^2/nb_{m_n}^2)EX^2I_{|X|\leq\tau b_{m_n}} \to 0.$ □

Most arguments in theorem 3.3 carry over to the stable case.

**Theorem 3.8.** For $(X_i)$ i.i.i. r.v. from a law satisfying (3.34) , $\tau$ as in the corollary 3.7 and $0 < p < \alpha$

$$b_{m_n}^{-1} \sum_{j=1}^{m_n} ( X_{nj}^{\omega} - n^{-1} \sum_{i=1}^{n'} X_i(\omega) I_{|X_i(\omega)| \leq \tau b_{m_n}} ) \xrightarrow{d_p} \theta \text{ in probability.}$$

**Proof.** Since we have weak convergence in probability it is enough to show that the p-moments are bounded in probability. The case $1 < \alpha < 2$ is similar to the gaussian case. I only will do the case $0 < \alpha \leq 1$. The technique of spliting the p-moments in different parts also work here. Consider first the case $0 < \tau < \infty$. Note that

$$E' |b_{m_n}^{-1} \sum_{j=1}^{m_n} ( X_{nj}^{\omega} - E' X_{nj}^{\omega} I_{|X_i(\omega)| \leq \tau b_{m_n}} )|^p \leq$$

$$2^{p-1} E' |b_{m_n}^{-1} \sum_{j=1}^{m_n} ( X_{nj}^{\omega} I_{|X_i(\omega)| \leq \tau b_{m_n}} - E' X_{nj}^{\omega} I_{|X_i(\omega)| \leq \tau b_{m_n}} )|^p +$$

$$2^{p-1} E' |b_{m_n}^{-1} \sum_{j=1}^{m_n} X_{nj}^{\omega} I_{|X_i(\omega)| \geq \tau b_{m_n}} |^p$$

As far the first part

$$[E' |b_{m_n}^{-1} \sum_{j=1}^{m_n} ( X_{nj}^{\omega} I_{|X_i(\omega)| \leq \tau b_{m_n}} - E' X_{nj}^{\omega} I_{|X_i(\omega)| \leq \tau b_{m_n}} )|^p]^{2/p} \leq$$

$$E' |b_{m_n}^{-1} \sum_{j=1}^{m_n} ( X_{nj}^{\omega} I_{|X_i(\omega)| \leq \tau b_{m_n}} - E' X_{nj}^{\omega} I_{|X_i(\omega)| \leq \tau b_{m_n}} )|^2 \leq (m_n / n b_{m_n}^2) \sum_{i=1}^{n} X_i^2 I_{|X_i| \leq b_{m_n}}$$

Also

$$E' |b_{m_n}^{-1} \sum_{j=1}^{m_n} X_{nj}^{\omega} I_{|X_i(\omega)| \geq \tau b_{m_n}} |^p \leq E' |b_{m_n}^{-p} \sum_{j=1}^{m_n} X_{nj}^{p} I_{|X_i(\omega)| \geq \tau b_{m_n}} | \leq (m_n / n b_{m_n}^p) \sum_{i=1}^{n} X_i^p I_{|x_i| \geq b_{m_n}}$$

So it is enough to show that $(m_n / n b_{m_n}^2) \sum_{i=1}^{n} X_i^2 I_{|X_i| \leq b_{m_n}}$ and $(m_n / n b_{m_n}^p) \sum_{i=1}^{n} X_i^p I_{|x_i| \geq b_{m_n}}$ are uniformly bounded.

This can be deduced from boundedness in $L_1$ using (3.42) and (3.43).

The general case follows from the facts that $n^{-1} \sum X_i I_{|X_i| \leq \tau b_{m_n}}$ converges in probability if $0 < \alpha < 1$. □

## 4. The a.s. bootstrap CLT of the mean.

Giné-Zinn (1989a) showed that if $\mathrm{E}X^2 = \infty$ and $m_n = n$ the CLT cannot be bootstrapped a.s. The same is true if $m_n \geq cn$ for some $c > 0$:

**Proposition 4.1.** If for some sequence $a_n \to \infty$, random variables $c_n(\omega)$ and random measures $\mu(\omega)$ non-degenerate with positive probability there exist $m_n \nearrow \infty$ such that $\inf m_n/n > 0$ for which

(4.1) $a_n^{-1} \sum_{j=1}^{m_n} X_{nj}^{\omega} - c_n(\omega) \xrightarrow{\mathcal{L}} \mu(\omega)$ a.s.

then $\mathrm{E}X^2 < \infty$.

**Proof.** By theorem 3.1 we can take $c_n(\omega) = (m_n/na_n) \sum_{i=1}^{n} X_i(\omega) I_{|X_i(\omega)| \leq a_n}$ and then $\mu(\omega) = \mathrm{N}(0, \sigma^2)$ a.s. for some $\sigma \in (0, \infty)$. Then (4.1) implies ( converse CLT) $(m_n/n) \sum_{i=1}^{n} I_{|X_i| > \delta a_n} \to 0$ a.s. for all $\delta > 0$. This and $\inf m_n/n > 0$ gives

(4.2) $\sum_{i=1}^{n} I_{|X_i| > \delta a_n} = 0$ eventually a.s.

hence, also

(4.2)' $\sum_{i=1}^{n} |X_i|^p I_{|X_i| > \delta a_n} = 0$ eventually a.s. for all p.

So, if $\mathrm{E}X^2 = \infty$ the truncated variance condition of the CLT for $X_{nj}^{\omega}$ becomes [recall (3.7)]

(4.3) $(m_n/na_n^2) \sum_{i=1}^{n} X_i^2 \to \sigma^2$ a.s.

If we let $b_n = a_n (n/m_n)^{1/2}$, then (4.3) gives,as in the last part of the proof of theorem 3.1, that X is in the domain of atraction of the normal law with norming constants $b_n$. Since $\mathrm{E}X^2 = \infty$ this implies in particular that $b_n \cong d_n$ ( in the sense that $b_n/d_n \to 1$) with $d_n = n^{1/2} L(n)$, where $L(n) \nearrow \infty$, Hence $d_n^2/n \nearrow \infty$, so that we can apply a result of Feller (e.g. Stout (1974) Theorem 3.2.5, p. 132) to conclude that the limit in (4.3) is either 0 or $+\infty$, a contradiction. Therefore, $\mathrm{E}X^2 < \infty$. □

We believe that Proposition 4.1 is not best possible. In fact, in view of the following result, it is possible that Poposition 4.1 holds true for all sequences $m_n \nearrow \infty$ such that $\liminf_{n\to\infty}(m_n L\,L\,n/n) > 0$, where $L\,r = \log(r \vee e)$, and $L\,L\,r = L(L\,r)$)

**Theorem 4.2.** If $\mathbf{E}X^2 = \infty$, there are a Lévy measure $\pi$, a real number $\sigma^2 \geq 0$ and random variables $c_n(\omega)$ so that

(4.4) $\sum_{j=1}^{m_n} X^{\omega}_{nj}/b_{m_n} - c_n(\omega) \xrightarrow{\mathcal{L}} \mathrm{N}(0,\sigma^2) * c_\tau Pos\,\pi$ a.s.

for each $\tau$ with $\pi\{-\tau,\tau\} = 0$.Assume also that $\sigma^2$ and $\pi$ are not simultaneously zero. If $\sup m_n/n < \infty$ then $A :=\inf m_n(L\ L\,n)/n = 0$.

**Proof.** The proof is by contradiction. Assume $A > 0$. We first consider the case $\pi \neq 0$. Take $\delta > 0$ such that $\pi\{-\delta,\delta\} > 0$ and $\pi\{|x| > \delta\} > 0$. By the converse CLT

(4.5) $(m_n/n)\ \sum_{i=1}^{n} I_{|X_i|>\delta b_{m_n}} \to \pi_\delta := \pi\{(\infty,-\delta)\cup(\delta,\infty)\}$ a.s.

Since $\sup m_n/n < \infty$, by the convergence of moments result in Acosta-Giné (1979) $m_n\mathbf{P}(|X| > a_n) \to \pi_\delta$. So (4.5) is equivalent to

(4.5)' $(m_n/n)\ \sum_{i=1}^{n}(I_{|X_i|>\delta b_{m_n}} - p_n) \to 0$ a.s.

Let $n_k = k^k$ and let $\bar{m}_k = m_{n_k}$, $\bar{a}_k = b_{\bar{m}_k}$ and $\bar{p}_k = p_{n_k}$. We claim

(4.6) $\sum_{k=1}^{\infty}\mathbf{P}\{(\bar{m}_k/n_k)\ \sum_{i=1}^{n_{k-1}}(I_{|X_i|>\bar{a}_k} - \bar{p}_k)| > K\} < \infty$ for all $K > 0$.

To estimate the probabilities in (4.6) we use Prokorov's exponential inequality (e.g. Stout (1974) Theorem 5.2.2, p. 262 ) which states that for $\xi_i$ indepent centered, $|\xi_i| \leq cs_n$ where $s_n^2 = \sum_{i=1}^{n}\mathbf{E}\xi_i^2$, $\mathbf{P}\{|\sum_{i=1}^{n}\mathbf{E}\xi_i/s_n| \geq \epsilon\} \leq 2\exp\{-(\epsilon/2c)\ \operatorname{arcsinh}\ [\epsilon c/2]\}$.

We take $\xi_i = I_{|X_i|>\bar{a}_k} - \bar{p}_k$, $\epsilon = (Kn_k/\bar{m}_k s_k)$ where $s_k^2 = n_{k-1}\bar{p}_k(1-\bar{p}_k)$, and $c = s_k^{-1}$. Since $\bar{m}_k\bar{p}_k \to \pi_1$ and $n_k/kn_{k-1} \to e$ we have for $\delta > 0$ and $k$ large enough,

$((1-\delta)eKk/2\pi_1) \leq \epsilon c/2 \leq ((1+\delta)eKk/2\pi)$

Hence we can replace arcsinh by $(1-\delta')\log$ in Prohorov's inequality. We obtain $\mathrm{P}\{(\bar{m}_k/n_k)\,|\sum_{i=1}^{n_k-1}(I_{|X_i|>\bar{a}_k}-\bar{p}_k)|>K\}\leq 2\exp\{-(Kn_k/2\bar{m}_k)(1-\delta')\log k\}\leq$ $2\exp(-2\log k)=2k^{-2}$ for k large enough, since $n_k/\bar{m}_k\rightarrow\infty$. (4.6) is proved. And (4.6) implies that for all $K>0$,

(4.7) $(\bar{m}_k/n_k)\,|\sum_{i=1}^{n_k-1}(I_{|X_i|>\bar{a}_k}-\bar{p}_k)|\leq K$ eventually a.s.

We show next that there exist $L>0$ such that

(4.8) $\sum_{k=1}^{\infty}\mathrm{P}\{(\bar{m}_k/n_k)\,|\sum_{i=n_{k-1}+1}^{n_k}(I_{|X_i|>\bar{a}_k}-\bar{p}_k)|>L\}=\infty$

To prove (4.8) we will invoke Kolmogorov's exponential minorization (Stout (1974), p. 262): for $\xi_i$ as above there are, for all $\gamma>0$, $\epsilon(\gamma)$ such that if $\epsilon\geq\epsilon(\gamma)$ and $\epsilon c\geq\pi(\gamma)$ then $\mathrm{P}\{|\sum_{i=1}^{n}\xi_i/s_n|\geq\epsilon\}\geq\exp\{-\epsilon^2(1+\gamma)/2\}$.

Now the variance of the sum is $\bar{s}_k^2:=(n_k-n_{k-1})\bar{p}_k(1-\bar{p}_k)$ so that, with $c=1/\bar{s}_k$ and $\epsilon=n_kL/\bar{m}_k\bar{s}_k$ we have, for large enough, $\epsilon c\leq 2L/\pi_1$ and $\epsilon\geq(L/2\pi_1^{1/2})\,(n_k/\bar{m}_k)^{1/2}\rightarrow\infty$. Hence, taking $\gamma=1$ we can apply Kolmogorov's inequality for $L\leq\pi(1)\,\pi_1/2$.

We then have, for k large enough

$\mathrm{P}\{(\bar{m}_k/n_k)\,|\sum_{i=n_{k-1}+1}^{n_k}(I_{|X_i|>\bar{a}_k}-\bar{p}_k)|>L\}\geq\exp[-2L^2\log k/\pi_1 A]$ which is the general term of a divergence series if $L^2\leq A\pi_1/2$.

Hence (4.8) holds for all $0<L\leq\pi(1)\pi_1/2\vee[A\pi_1/2]^{1/2}$. But (4.8) implies ( by disjointness of the intervals $(n_{k-1},n_k]$ and Borel-Cantelli) that

(4.9) $(\bar{m}_k/n_k)\,|\sum_{i=n_{k-1}+1}^{n_k}(I_{|X_i|>\bar{a}_k}-\bar{p}_k)|>L$ a.s. infinitely often.

(4.7) for $K=L/2$ and (4.9) show that a.s. $(\bar{m}_k/n_k)\ \sum_{i=n_{k-1}+1}^{n_k}(I_{|X_i|>\bar{a}_k}-\bar{p}_k)|>L/2>0$ infinitely often, therefore (4.5)' does not hold.

If $\pi=0$, then $\sigma^2\neq 0$. By the converse CLT and Lemma 2.1

(4.10) $(m_n/na_{m_n}^2)\sum_{i=1}^{n}X_i^2I_{|X_i|\leq a_{m_n}}\rightarrow\sigma^2$ a.s.

Again by Acosta-Giné (1979) $(m_n/a_{m_n}^2)U(a_{m_n})\rightarrow\sigma^2$. But a proof completely analogous to the above one shows that for some $L > 0$ $(\bar{m}_k/n_k\bar{a}_k^2)\ |\sum_{i=1}^{n_k}(X_i^2 I_{|X_i|\le \bar{a}_k} - U(\bar{a}_k))| > L/2 > 0$ a.s.infinitely often, which contradicts (4.10). □

Finally we show that the a.s. bootstrap CLT for X in the domain of atraction of any stable law always holds if $m_n(L\,L\,n)/n \to 0$, at least under regularity of $\{m_n\}$ ($m_n/m_{2n} \ge c$ for some $c > 0$ and $m_n \nearrow \infty$). Theorem 4.2 shows that the result is sharp. Thus, the newt theorem improves Athreya's result (Corollary 3.6 above) for these sequences.

**Theorem 4.3.** Let $\theta$ be a nondegenerate p-stable random variable, $0 < p \le 2$, and let X be a random variable in its domain of atraction, concretely, let **X** satisfy

(4.11) $\sum_{i=1}^{n}(X_i - \mathbf{E}XI_{|X|\le \tau b_n}))/b_n \xrightarrow{\mathcal{L}} \theta$

with $b_n \nearrow \infty$. Let $\{m_n\}$ be a sequence of positive integers regular in the sense that $m_n \nearrow \infty$ and $m_n/m_{2n} \ge c$ for some $c > 0$ and all $n \in \mathbf{N}$ and such that

(4.12) $m_n(L\,L\,n)/n \to 0$

then

(4.13) $\sum_{j=1}^{m_n}(\,X_{nj}^{\omega} - n^{-1}\sum_{i=1}^{n} X_i(\omega)I_{|X_i(\omega)|\le b_{m_n}})/b_{m_n} \xrightarrow{\mathcal{L}} \theta$ a.s.

**Proof.** By Bickel and Freedman's theorem, only the case $\mathbf{E}X^2 = \infty$ requires proof. Let, for $\lambda > 0$, $\delta > 0$, $\pi_\lambda = \lim_{n\to\infty} n\mathbf{P}\{X > \lambda b_n\}$ $\quad \pi_{-\lambda} = \lim_{n\to\infty} n\mathbf{P}\{X < -\lambda b_n\}\sigma_\delta^2 = \lim_{n\to\infty}(n/b_n^2)U(\delta b_n)$ and $\sigma^2 = \lim_{\delta\to 0}\sigma_\delta^2$.

As usual, it suffices to show that $(m_n/n)\ \sum_{i=1}^{n} I_{X_i > \delta b_{m_n}} \to \pi_\lambda$ a.s. $(m_n/n)\sum_{i=1}^{n} I_{X_i < -\delta b_{m_n}} \to \pi_{-\lambda}$ a.s. and $\lim_{\delta\to 0}\lim_{n\to\infty}(m_n/nb_{m_n^2})\sum_{i=1}^{n} X_i^2 I_{|X_i|\le \delta b_{m_n}} = \sigma^2$ a.s. ($\pi_\lambda, \pi_{-\lambda} = 0$ if $p = 2$; $\sigma^2 = 0$ if $p < 2$; for $p = 2$, $\lim_{\delta\to 0}$ is redundant). Letting $p_{n,\lambda} = \mathbf{P}\{X > \lambda b_{m_n}\}$ and $p_{n,-\lambda} = \mathbf{P}\{X > -\lambda b_{m_n}\}$ ,these limits are equivalent to

(4.14) $(m_n/n) \sum_{i=1}^{n}(I_{X_i>\delta b_{m_n}} - p_{n,\lambda}) \rightarrow 0$ a.s.

(4.15) $(m_n/n) \sum_{i=1}^{n}(I_{X_i<-\delta b_{m_n}} - p_{n,-\lambda}) \rightarrow 0$ a.s. and

(4.16) $\lim_{\delta\rightarrow 0} \lim_{n\rightarrow\infty}(m_n/nb_{m_n^2} \sum_{i=1}^{n}[X_i^2 I_{|X_i|\leq\delta b_{m_n}} - U(\delta b_{m_n})] = 0$ a.s.

The three limits can be proved using exactly the same technique. So, we give the details of the proof only for (4.14). By Borel-Cantelli, (4.14) will follow if we show

(4.17) $\sum_{n=1}^{\infty} \mathbf{P}\{max_{2^{n-1}<k\leq 2^n}(m_k/k) \, |\sum_{i=1}^{k}(I_{X_i>\lambda b_{m_k}} - p_{k,\lambda})| > \epsilon\} < \infty.$

for all $\epsilon > 0$. To prove this we symmetrize, apply Lévy's maximal inequality, and then use an exponential inequality to estimate the resülting probability. For the symmetrization we use an idea of Hoffmann-Jorgensen (1974) (proof of Corollary 3.4). Let us consider the following $l_{2^{n-1}}^{\infty}$ valued vectors $v_i = ((m_k/k)I_{X_i>\lambda b_{m_k}} : 2^{n-1} < k \leq 2^n), \;\; i = 1, \ldots, 2^{n-1}+1$

$v_i = (0, \ldots, 0, (m_k/k)I_{X_i>\lambda b_{m_k}} : i \leq k \leq 2^n) \quad i = 2^{n-1}+2, \ldots, 2^n$. Then

(4.18) $\|\sum_{i=1}^{2^n}(v_i - \mathbf{E}v_i)\| = max_{2^{n-1}<k\leq 2^n}(m_k/k) \, |\sum_{i=1}^{k}(I_{|X_i|>\lambda b_{m_k}}) - p_{k,\lambda})|.$

If $(v)^r$ denotes the r-th coordinate of $v \in l_{2^{n-1}}$, then

$\max_r \mathbf{P}\{|\sum_{i=1}^{2^n}(v_i - \mathbf{E}v_i)^r| > \epsilon\} \leq \epsilon^{-1} max_{2^{n-1}<k\leq 2^n}(m_k/k) \, |\sum_{i=1}^{k}(I_{|X_i|>\lambda b_{m_k}} - p_{k,\lambda}| \leq$
$\epsilon^{-1} max_{2^{n-1}<k\leq 2^n}(m_k/k)(kp_{k,\lambda})^{1/2} = \epsilon^{-1} max_{2^{n-1}<k\leq 2^n}(m_k/k)^{1/2}(m_k p_{k,\lambda})^{1/2} \rightarrow 0$ since $m_k p_{k,\lambda} \rightarrow \pi_\lambda$ and $m_k/k \rightarrow 0$. Therefore we can apply to $\sum_{i=1}^{2^n}(v_i - \mathbf{E}v_i)$ e. g. the symmetrization lemmas in Giné-Zinn (1984) (lemma 2.5) to obtain that for each $\epsilon > 0$ there is $n(\epsilon) < \infty$ such that for $n > n(\epsilon)$,

(4.19) $\mathbf{P}\{\|\sum_{i=1}^{2^n}(v_i - \mathbf{E}v_i)\| > \epsilon\} \leq 2\mathbf{P}\{\|\sum_{i=1}^{2^n}(v_i - \mathbf{E}v_i)\| > \epsilon/3\}$

where $\{\epsilon_i\}$ is an independent sequence of i.i.d. random variables with $\mathbf{P}\{\epsilon_i = 1\} = \mathbf{P}\{\epsilon_i = -1\} = 1/2$, independent of $\{v_i\}$. By (4.18) and (4.19), the proof of (4.17) reduces to showing

(4.20) $\sum_{n=1}^{\infty} \mathbf{P}\{max_{2^{n-1}<k\leq 2^n(m_k/k)} \, |\sum_{i=1}^{k} \epsilon_i I_{X_i>\lambda b_{m_k}}| > \epsilon\} < \infty$ for all $\epsilon > 0$.

In order to apply P. Lévy's maximal inequality we write

$\mathbf{P}\{max_{2^{n-1}<k\leq 2^n(m_k/k)} \, |\sum_{i=1}^{k} \epsilon_i I_{X_i>\lambda b_{m_k}}| > \epsilon\} \leq$

$\mathbf{P}\{(m_{2^{n-1}}/2^{n-1}) max_{2^{n-1}<k\leq 2^n(m_k/k)} |\sum_{i=1}^{2^n} \epsilon_i I_{X_i>\lambda b_{m_k}}| > c\epsilon/2\}+$

$\mathbf{P}\{(m_{2^{n-1}}/2^{n-1}) max_{2^{n-1}<k\leq 2^n(m_k/k)} |\sum_{i=k+1}^{2^n} \epsilon_i I_{X_i>\lambda b_{m_k}}| > c\epsilon/2\}.$

and notice that the sets of indices $A_k = \{i \leq 2^n : X_i > \lambda b_{m_k}\}$ $k = 2^{n-1}, \ldots, 2^n$ and $B_k = \{i \leq 2^n : k+1 \leq i \leq 2^n, X_i > \lambda b_{m_k}\}$ $k = 2^{n-1}, \ldots, 2^n$ are both decreasing as k increases since $b_{m_k}$ increases with $k$. Therefore, for $X_1 \ldots, X_{2^n}$ fixed, the above maxima are actually maxima of partial sums respectively of $\sum_{i\in A_{2^n-1}} \epsilon_i$ and $\sum_{i\in B_{2^n-1}} \epsilon_i$, suitable ordered. So we can apply Lévy's maximal inequality conditionally on the $X_i$'s and then integrate with respect to the $X_i$'s. Taking into account that $B_{2^{n-1}} \subset A_{2^{n-1}}$, so that we can apply Lévy's inequality twice to the second probability, we obtain

(4.21) $\mathbf{P}\{max_{2^{n-1}<k\leq 2^n(m_k/k)} |\sum_{i=1}^{k} \epsilon_i I_{X_i>\lambda b_{m_k}}| > \epsilon\} \leq$

$6\ \mathbf{P}\{(m_{2^{n-1}}/2^{n-1}) |\sum_{i=1}^{2^n} \epsilon_i I_{X_i>\lambda b_{m_{2^{n-1}}}}| > c\epsilon/2\}.$

Since $m_{2^{n-1}} p_{2^{n-1},\lambda} \rightarrow \pi_\lambda$, $2^{n-1}/m_{2^{n-1}} \geq K_n \log n$ and $K_n \rightarrow \infty$, Prohorov's inequality ( stated in the proof of Theorem 4.2) apllied to the last probality in (4.21) shows that for $n$ large,

(4.22) $\mathbf{P}\{max_{2^{n-1}<k\leq 2^n(m_k/k)} |\sum_{i=1}^{k} \epsilon_i I_{X_i>\lambda b_{m_k}}| > \epsilon\} \leq$

$12 \exp\{-(c\epsilon 2^{n-1}/8m_{2^{n-1}}) \operatorname{arcsinh}[c\epsilon/4m_{2^{n-1}}p_{2^{n-1},\lambda}]\} \leq 12 \exp\{-dK_n \log n\}.$

where $d$ is a fixed constant. (4.22) is the general term of a convergent series and therefore (4.20) converges. This proves (4.14). □

Now I present some results about the convergence of the Mallows distance in the case of the normal domain of atracttion to a normal random variable.

**Theorem 4.4.** Given $m_n \nearrow \infty$

(a) $EX^2 < \infty$ if and only if for some $\sigma < \infty$

$m_n^{-1/2} \sum_{i=1}^{m_n} (X_{n\,i} - \bar{X}_n) \xrightarrow{w} \mathrm{N}(0, \sigma)$ a.s.

In this case $m_n^{-1/2}\sum_{i=1}^{m_n}(X_{n\,i} - \bar{X}_n) \xrightarrow{d_2} \mathbf{N}(0,\sigma)$ a.s

(b) For $2 \leq p$, $EX^2 < \infty$ and $\sum_{n=1}^{\infty} P\{|X| \geq m_n^{1/2-1/p} n^{1/p}\} < \infty$ together are equivalent to $m_n^{-1/2}\sum_{i=1}^{m_n}(X_{n\,i} - \bar{X}_n) \xrightarrow{d_p} \mathbf{N}(0,\sigma)$ a.s

(c) $\sum_{n=1}^{\infty} P\{|X| \geq \epsilon m_n^{1/2}\} < \infty$ for all $\epsilon > 0$ and $EX^2 < \infty$ implies that for any $t \in \mathbf{R}$ $\exp[t m_n^{-1/2}\sum_{i=1}^{m_n}(X_{n\,i} - \bar{X}_n)] \rightarrow E\exp[t\sigma g]$ where g is a $\mathbf{N}(0,\sigma)$ r.v. and $\sigma^2 = Var(X)$

**Proof of a.** Only I need to show that if $m_n^{-1/2}\sum_{i=1}^{m_n}(X_{n\,i} - \bar{X}_n) \xrightarrow{w} \mathbf{N}(0,\sigma)$ a.s. then $EX^2 < \infty$ . By the converse of the CLT $n^{-1}\sum_{i=1}^{n} X_i^2 I_{|X_i| \leq m_n} - (n^{-1}\sum_{i=1}^{n} X_i I_{|X_i| \leq m_n})^2 \rightarrow \sigma^2$ a.s. By Lemma 2.1 $n^{-1}\sum_{i=1}^{n} X_i^2 I_{|X_i| \leq m_n^{1/2}} \rightarrow \sigma^2$ a.s.

Take $0 < c < \infty$, then $n^{-1}\sum_{i=1}^{n} X_i^2 I_{|X_i| \leq c} \leq n^{-1}\sum_{i=1}^{n} X_i^2 I_{|X_i| \leq m_n}$. So $EX^2_{|X| \leq c} \leq \sigma^2$. Hence $EX^2 < \infty$

**Proof of b.** Suppose that $2 < p$, $EX^2 < \infty$ and $\sum_{n=1}^{\infty} P\{|X| \geq m_n^{1/2-1/p} n^{1/p}\} < \infty$. By problem 13 in page 69, Araujo-Giné (1980) we need to show

$\lim_{t \rightarrow 0}\sup_n m_n^{1-p/2} n^{-1}\sum_{i=1}^{n} |X_i|^p I_{|X_i| \geq m_n^{1/2} t} = 0$ a.s. and

$\lim_{n \rightarrow \infty} m_n^{1/2} n^{-1}\sum_{i=1}^{n} |X_i| I_{|X_i| \geq m_n^{1/2}} = 0$ a.s.

Note that by the result by Feller that we used in Theorem 4.1 if $\sum_{n=1}^{\infty} P\{|X| \geq m_n^{1/2-1/p} n^{1/p}\} < \infty$ and $m_n \nearrow \infty$ then $m_n^{1-p/2} n^{-1}\sum_{i=1}^{n} |X_i|^p \rightarrow 0$ a.s. The rest is routine.

**Proof of c** As in the proof of the theorem 3.3

$E'\exp(t\, m_n^{-1/2}\sum_{i=1}^{m_n}(X_{n\,i}^{\omega} - \bar{X}_n(\omega))) \leq [n^{-1}\sum_{i=1}^{n} \exp(2t^2 X_i^2 a_n^{-2})]^{m_n}$

By the Borell-Cantelly lemma,the hypothesis is equivalent that $\max_{i \leq n}|X_i|^2/m_n \rightarrow 0$ a.s.

So $m_n \log[n^{-1}\sum_{i=1}^{n} \exp(2t^2 X_i^2 m_n^{-1})] \cong n^{-1}\sum_{i=1}^{n} 2t^2 X_i^2 \rightarrow 2t^2 EX^2$ a.s. □

## 5. Closing Remark.

The problem of how to solve the failure of the bootstrap of the mean in the stable case is a work in progress. For example in Wu-Carlstein-Cambanis (1989) there is another alternative:

Supose $X_1, \ldots, X_n$ are i.i.d. random variables. Partition the data set into $\mathrm{l}_n$ blocks, each having $k_n$ observations $n = k_n \, \mathrm{l}_n$. Instead of bootstrapping from the entire sample, they apply the booootstrap algorithm within each block $j = 1, \ldots, \mathrm{l}_n$ to obtain a block version version of the sample mean. Then they take the average of the $\mathrm{l}_n$ bootstrap distribution that they have obtained.

# CHAPTER III

# THE BOOTSTRAP OF U-STATISTICS AND V-STATISTICS

## 1. Definition of U-statistics and V-statistics.

The aim of this chapter is to describe the paper of Arcones and Giné (1990a) about the bootstrap of U and V statistics. Let's start by introducing U-statistics and V-statistics. Let $(X_i)$ be i.i.d.r.v.'s with values in a probability space $(S, \mathcal{S}, \mathbf{P})$. The U and V statistics based on h and P are defined respectively as

(1.1) $U_m^n(h, \mathbf{P}) = \binom{n}{m}^{-1} \sum_{1 \le i_1 < \cdots < i_m \le n} h(X_{i_1} \ldots, X_{i_m})$.

(1.2) $V_m^n(h, \mathbf{P}) = n^{-m} \sum_{i_1, \ldots, i_m = 1}^{n} h(X_{i_1}, \ldots, X_{i_m})$.

The first definition was introduced in Hoeffding (1948a) and the second one in von Mises (1947).

These two classes of estimators are important in first place because some usual estimators are in these classes. For instance the smaple variance is the V-statistic over the kernel $h(x_1, x_2) = (x_1 - x_2)^2/2$.

Another reason why these estimators are important is because, in analogy to Taylor series of differentiable functions, it is possible to define differentiability for estimators. Since the Taylor expansion the k-term is a k-linear form by Frechet theorem, unless for tame linear continuous functionals, we can expect that under regularity conditions, the k-term in a expansion over the probability measures is given by $\int h_k(x_1, \ldots, x_k) d\mathbf{P}(x_1) \cdots d\mathbf{P}(x_k)$. But

$\int h_k(x_1, \ldots, x_k) d\mathbf{P}_n(x_1) \cdots d\mathbf{P}_n(x_k) = n^{-k} \sum_{i_1 \ldots i_k = 1}^{n} h_k(X_{i_1}, \ldots, X_{i_k}) = V_m^n(h, P)$.

Therefore an estimator is differentiable in a certain sense if

(1.3) $T(X_1, \ldots, X_n) = h_0 + \sum_{i=1}^{n} h_1(X_i) + \sum_{i,j=1}^{n} h_2(X_i, X_j) + \cdots + \sum_{i_1 \ldots i_k = 1}^{n} h_k(X_{i_1}, \ldots, X_{i_k}) +$
$R_k(X_1, \ldots, X_n)$

where $R_k(X_1, \ldots, X_n) \to 0$ in a convenient way. This is the so called von Mises expansion.

Given that a V-statistic can be expanded in U-statistics we can also have a U-statistics expansion

(1.4) $T(X_1, \ldots, X_n) = \sum_{j=0}^{k} U_j^n(h_j) + R_k(X_1, \ldots, X_n)$.

For more information about U-statistics we refer to Chapter 5 in Serfling (1980).

## 2. Prelimary results.

A function $h : \mathbf{S}^m \to \mathbf{R}$ is measurable symmetric if $h(x_1, \ldots, x_m) = h(x_{\sigma(1)}, \ldots, x_{\sigma(m)})$ for all $(x_1, \ldots, x_m) \in \mathbf{S}^m$ and all permutations $\sigma$ of $N_m = \{1, \ldots, m\}$.

It is convenient to define the following two linear operators on (symmetric) functions $h : \mathbf{S}^m \to \mathbf{R}$. For $x_1, \ldots, x_m \in \mathbf{S}$ $n \in \mathbf{N}$

(2.1) $\sigma_m^n h(x_1, \ldots, x_n) = \sum_{1 \le i_1 < \cdots < i_m \le n} h(x_{i_1} \ldots, x_{i_m})$ for $1 \le m \le n$

$\sigma_m^n h(x_1, \ldots, x_n) = 0$ if $m > n$

$\sigma_0^n(c) = c$ for $c \in \mathbf{S}$.

The projection operator $\pi_{k,m}^{\mathbf{P}}$ is defined on functions h of m variables in $L_1(\mathbf{S}^m, \mathcal{S}^{\ m}, \mathbf{P}^m)$ and takes values in $L_1(\mathbf{S}^k, \mathcal{S}^{\ k}, P^k)$ for $0 \le k \le m$ as follows for $(x_1, \ldots, x_m) \in \mathbf{S}^m$ $m \ne 0$

(2.2) $\pi_{k,m}^{\mathbf{P}}(h)(x_1 \ldots, x_k) = (\delta_{x_1} - \mathbf{P}) \cdots (\delta_{x_k} - P)\mathbf{P}^{m-k} h$

$\pi_{0,0}^{\mathbf{P}}(c) = c$ for $c \in \mathbf{S}$

where for a measure $\mathbf{Q}_i$ on $\mathbf{S}$ $\mathbf{Q}_1 \cdots \mathbf{Q}_m h = \int_{\mathbf{S}^n} h(x_1, \ldots, x_m)\ dQ_1(x_1) \cdots, d\mathbf{Q}_m(x_m)$. If a function is not necessarily symmetric, we will write $S_m h$ for its symmetrization, i.e.

(2.3) $(S_m h)(x_1, \cdots, x_m) = (m!)^{-1} \sum h(x_{\sigma(1)}, \ldots, x_{\sigma(m)})$

where the sum extends over all permutations $\sigma$ of $N_m$. Important types of functions in this theory are the **P**-canonical. A function h is **P**-canonical if h is symmetric and $\mathbf{P}h(x_1, \cdots, x_{m-1}, \cdot) = \delta_{x_1} \cdots \delta_{x_{m-1}} \mathbf{P}h = 0$ for all $x_1, \ldots, x_{m-1} \in \mathbf{S}$. Note that if h is symmmetric then $\pi_{k,m}^{\mathbf{P}} h$ is P canonical for $1 \le k \le m$.

**Lemma 2.1.** $\sigma$ and $\pi$ are linear. Given $h : \mathbf{S}^m \to \mathbf{R}$ is symmetric

(a) If h is P-canonical then $\mathbf{E}(\sigma_m^n h(\mathbf{X}_1, \ldots, \mathbf{X}_n))^2 = \binom{n}{m} \mathbf{E} h^2(\mathbf{X}_1, \ldots, \mathbf{X}_m)$ for $1 \le m \le n$

(b) $\mathbf{E}(\pi_{k,m}^P h(\mathbf{X}_1, \ldots, \mathbf{X}_k))^2 \le \mathbf{E} h^2(\mathbf{X}_1, \ldots, \mathbf{X}_m)$ for $1 \le k \le m$

(c) $\sigma_m^n \, \sigma_k^m = \binom{n}{m} \binom{m}{k} \binom{n}{k}^{-1} \sigma_k^n$ for $0 \le k \le m \le n$

(d) $\pi_{j,k}^Q \pi_{k,m}^P = \pi_{j,k}^Q P^{m-k} h$ for $0 \le k \le m$

(e) If $h : \mathbf{S}^r \to \mathbf{R}$ is **P**-canonical $0 \le r, s \le m$ then $\pi_{s,m}^P \sigma_r^m h = h$ if $r = s$ $= 0$ if $r \ne s$.

**Proof.** (a) follows from the fact that if h is P-canonical then $\mathbf{E} h(X_{i_1} \ldots, X_{i_m}) h(X_{j_1} \ldots, X_{j_m}) = 0$ unless the sets $\{i_1, \ldots, i_m\}$ and $\{j_1, \ldots, j_m\}$ are equal. (b) follows applying conditionally that $\mathbf{E}(X - \mathbf{E}X)^2 \le \mathbf{E}X^2$. (c) follows from the fact in $\sigma_m^n \, \sigma_k^m h$ we have the same number of times every summand in $\sigma_k^n h$. (d) follows from Fubini's theorem. (e) is trivial. □

Using the $\sigma - \pi$ notation we obviously have that for h symmetric

(2.4) $h(x_1, \ldots, x_m) = \delta_{x_1} \cdots \delta_{x_m} h = (\delta_{x_1} - P + P) \cdots (\delta_{x_m} - P + P) h = \sum_{k=0}^m \sigma_k^m \, \pi_{k,m}^{\mathbf{P}} h(x_1, \ldots, x_m)$.

Since $U_m^n(h, P) = \binom{n}{m}^{-1} \sigma_m^n h(X_{i_1} \ldots, X_{i_m})$ it follows that

(2.5) $U_m^n(h, P) = \sum_{k=0}^m \binom{m}{k} \binom{n}{k}^{-1} \sigma_k^n \, \pi_{k,m}^P h(X_1, \ldots, \mathbf{X}_m) = \sum_{k=0}^m \binom{m}{k} U_k^n(\pi_{k,m}^P h, P)$.

This is the Hoeffding decomposition of $U_m^n$ into sum of U-statistics of P-canonical functions.

The symmetric function $h(x_1, \ldots, x_m)$, or the U-statistic $U_m^n(h, P)$ is of P-canonical order k $1 \le k \le m$ if $\pi_{1,m}^P = \cdots = \pi_{k-1,m}^P = 0$ and $\pi_{k,m}^{\mathbf{P}} \ne 0$. Unless h is a constant it has an canonical order and then its Hoeffding decomposition is

(2.6) $U_m^n(h, P) - \mathbf{E} U_m^n(h, P) = \sum_{r=k}^m \binom{m}{r} \binom{n}{r}^{-1} \sigma_r^n \, \pi_{r,m}^P h(X_1, \ldots, X_m) = \sum_{r=k}^m \binom{m}{r} U_r^n(\pi_{r,m}^P h, P)$.

The order of $U_m^n(h, P)$ is m. Of course if h is P-canonical its canonical order is m. Note that by lemma 2.1

(2.7) $\|n^{k/2}[U_m^n(h,P) - \mathbf{E}U_m^n(h,P) - \binom{m}{k}U_k^n(\pi_{k,m}^P h, P)]\|_2^2 \le \sum_{r=k+1}^m n^k \binom{m}{r}^2 \binom{n}{r}^{-1} \|h\|_2^2 \to 0.$

We also need the following lemma

**Lemma 2.2.** Suppose $(S, d)$ is a metric space. Let $\{X_{n\,\tau};\ n \in \mathbf{N}\ \ \tau > 0\}$ be a family of random variables with values in S. Assume that $X_{n\,\tau} \to_w Y_\tau$ as $n \to \infty$ for each $\tau > 0$ and for each $\epsilon > 0$ $\lim_{\tau \to 0} \limsup_{n \to \infty} P\{d(X_{n\,\tau}, X_n) > \epsilon\} = 0$ and $\lim_{\tau \to 0} P\{d(Y_\tau, Y) > \epsilon\} = 0$ then $X_n \to_w Y$.

**Proof.** Let $h : S \to \mathbf{R}$ be a function in $BL_1$. From

$|\mathbf{E}H(X_n) - \mathbf{E}H(Y)| \le |\mathbf{E}H(X_n) - \mathbf{E}H(X_{n,\tau})| + |\mathbf{E}H(X_{n,\tau}) - \mathbf{E}H(Y_\tau)| + |\mathbf{E}H(Y_\tau) - \mathbf{E}H(Y)| \le$

$\epsilon + 2P\{d(X_n, X_{n,\tau}) > \epsilon\} + |\mathbf{E}H(X_{n,\tau}) - \mathbf{E}H(Y_\tau)| + \epsilon + 2\mathrm{P}\{d(Y_\tau, Y) > \epsilon\}$

taking limits we get the result. □

In the proof of the bootstrap CLT we will require the law of large numbers and the central limit theorem for U-statistics. The first due to Hoeffding (1961) ( see also Berk (1966)) is as follows:

(2.8) $\mathbf{E}|h(X_1, \ldots, X_m)| < \infty$ implies that $U_m^n(h,P) \to \mathbf{E}U_m^n(h,P)$ a.s.

We will also require an easy complement of this theorem Giné-Zinn (1989). Let $h : \mathbf{S}^m \to \mathbf{R}$ be a symmetric function; then if $r > m$

(2.9) $\mathbf{E}|h(X_1, \ldots, X_m)|^{m/r} < \infty$ implies that $n^{-r} \sum_{1 \le i_1 < \cdots < i_m \le n} h(X_{i_1} \ldots, X_{i_m}) \to 0$ a.s.

A consequence for V-statistics is that if for each $i_1, \ldots, i_m$, $\{i_1, \ldots, i_m\} \subset \{1, \ldots, n\}$

(2.10) $\mathbf{E}|h(X_{i_1}, \ldots, X_{i_m})|^{\#\{i_1, \ldots, i_m\}/m} < \infty$ then

$$n^{-m} \sum_{i_1, \ldots, i_m = 1}^n h(X_{i_1}, \ldots, X_{i_m}) \to \mathbf{E}h(X_1, \ldots, X_m) \text{ a.s.}$$

### 3. The CLT for U-statistics.

First I will sketch a modification of Bretagnolle's proof of the CLT for U-statistics borrowing some notation from Dynkin-Mandelbaum (1983). For $\phi : \mathbf{S} \to \mathbf{R}$ measurable and bounded with $P\phi = 0$ let $h_k^\phi(x_1, \dots, x_k) = \phi(x_1) \cdots \phi(x_k)$.

Remember Newton's identities if $p_1 = \sum_{i=1}^n x_i \quad p_2 = \sum_{i<j}^n x_i x_j \qquad \cdots \quad p_n = x_1 \cdots x_n$ and $s_j = \sum_1^n x_i^j$ then for $k < n \quad s_k - p_1 s_{k-1} + p_2 s_{k-2} + \cdots + (-1)^{k-1} p_{k-1} s_1 + (-1)^k k p_k = 0$.

So $p_k = P_k(s_1, \dots, s_k)$ where $P_k$ is k-degree k-variable polinomial. So this implies

(3.1) $\quad n^{-k/2} \sigma_k^n h_k^\phi(x_1, \dots, x_k) = P_k(n^{-1/2} \sum_1^n \phi(x_i), n^{-1} \sum_1^n \phi^2(x_i), \dots, n^{-k} \sum_1^n \phi^{k/2}(x_i))$.

Let $\{G_{\mathbf{P}}(\phi) : \phi \in L_2(\mathrm{S}, \mathcal{S}, P)$ and $P\phi = 0\}$ be the isonormal Gaussian process i.e. $G_P(\phi)$ is $N(0, P\phi^2)$ and $Cov(G_{\mathbf{P}}(\phi), G_P(\psi)) = Cov(\phi, \psi)$ for all $\psi, \psi \in L_2(\mathrm{S}, \mathcal{S}, P)$. Note that $G_P$ is a linear operator.

Then the central limit theorem and the law of large numbers give

(3.2) $\mathcal{L}\ (n^{-k/2} \sigma_k^n h_k^{\phi_j}(\mathrm{X}_1, \dots, \mathrm{X}_n) : j = 1 \dots J) \to_w \mathcal{L}\ (P_k(G_P(\phi_j), \sigma_j^2, 0, \dots, 0) : j = 1, \dots, J)$

where $\sigma_\phi^2 = P\phi^2$.

In the case that $P\phi^2 = 1 \quad (P_k(G_P(\phi), 1, 0 \dots, 0)/k!)_k$ are orthonormal polinomials and the index is equal to its degree. Thus $P_k(G_P(\phi), 1, 0 \dots, 0) = H_k(G_P(\phi))/(k!)^{1/2}$ where $H_k$ is the k-th Hermite polinomial. By homogeinity $P_k(G_P(\phi), P(\phi^2), 0 \dots, 0) = (P\phi^2)^{k/2} H_k(G_P(\phi)/(P\phi^2)^{1/2})/(k!)^{1/2}$.

Let $(\Omega, \mathcal{S}, Q)$ be the probability space supporting the $G_{\mathbf{P}}(\phi)$. Define L as the linear span of $\{h_k^\phi ; \phi \in L_\infty(\mathbf{S}, \mathcal{S}, \mathbf{P})$ and $P\phi = 0\}$. Define also the following operator on L

(3.3) $\quad I(\sum_{j=1}^J c_j h_k^{\phi_j}) = \sum_{j=1}^J c_j \sigma_{\phi_j}^k H_k(G_P(\phi_j/\sigma_j))$

where $\sigma_j^2 = Var\phi_j$.

From 3.2. and the convergence of moments it follows that I is a linear isometry so it extends to the closure of L in $L_2(\mathrm{S}^k, \mathcal{S}^k, P^k)$. Call I this extension. Now (3.2) implies that

(3.4) $\mathcal{L}\,(n^{-k/2}\sigma_k^n h) \to_w \mathcal{L}\,(I(h))$ for any $h \in$ L.

To finish the proof we would like to prove (3.4) for h canonical. By real analisis we know that if h is a function of m variables with $Eh^2(X_1, \ldots, X_m) < \infty$, then h can be aproximated in $L_2(\mathbf{S}^k, \mathcal{S}^{\,k}, P^k)$ by functions of form $g = \sum_{i=1}^{s} c_i I_{A_{i1}}(x_1) \cdots I_{A_{ik}}(x_k)$ where $A_{ij} \in \mathcal{S}$ . Suppose besides that h is canonical then by lemma 2.1-b $\|h - \pi_{k\,k} g\|_2 = \|\pi_{k,k}(h-g)\|_2 \leq \|h-g\|_2$. Call $\phi_{ij}(x_j) = I_{A_{ij}}(x_j) - P(A_{ij})$ Then $\mathbf{P}\psi_{ij} = 0$ and $\pi_{k\,k}(g) = \sum_{i=1}^{s} c_i \phi_{i,1}(x_1) \ldots \phi_{i\,k}(x_k)$.

Since the operator $S_k : L_2(\mathbf{S}^k, \mathcal{S}^{\,k}, P^k) \to L_2(\mathbf{S}^k, \mathcal{S}^{\,k}, P^k)$ is a projection $\|h - S_k \pi_{k\,k} g\|_2 \leq \|h-g\|_2$. By polarization

$$\sum_\sigma \phi_{\sigma(1)}(x_1) \ldots \phi_{\sigma(k)}(x_k) = h_k^{\phi_1 + \cdots + \phi_k} - \sum_{i_1 \cdots i_{k-1}} h_k^{\phi_{i_1} + \cdots + \phi_{i_{k-1}}} + \cdots + (-1)^{k+1} \sum_{i=1}^{k} h_k^{\phi_i}.$$

So $S_k \pi_{k\,k} g = \sum_{j=1}^{J} c_j \psi_j(x_1) \cdots \psi_j(x_k)$ for some simple functions $\psi_i$ with $P(\psi_i) = 0$. This means that $S_k \pi_{k\,k} g$ is in L and h is in the closure of L. Given that the canonical funtions form the invariant space of the linear functionals $S_k$ and $\pi_{k\,k}$

(3.5) the closure of L consits of the canonical functions in $L_2(\mathbf{S}^k, \mathcal{S}^{\,k}, P^k)$.

So if h is canonical there are functions $g_\tau$ in L so that $\|h - g_\tau\|_2 \to 0$. By lemma 2.1 $\|n^{-k/2}\sigma_k^n h - n^{-k/2}\sigma_k^n g_\tau\|_2 \leq n^{-k}\binom{n}{k}\|h - g_\tau\|_2$ and $\|I(h) - I(g_\tau)\|_2 = \|h - g_\tau\|_2$. So by lemma 2.2. $\mathcal{L}\,(n^{-k/2}\sigma_k^n\, h) \to_w \mathcal{L}\,(I(h))$ for any canonical $h \in L_2(\mathbf{S}^k, \mathcal{S}^{\,k}, \mathbf{P}^k)$.

Combining the CLT just proved with (2.7) we obtain the CLT for U-statistics:

**Theorem 3.1.** If $h \in L_2(\mathbf{S}^m, \mathcal{S}^{\,m}, P^m)$ has canonical order k then the sequence

(3.6) $\{n^{k/2}(U_m^n(h,P) - \mathbf{E}U_m^n(h,P))\}$ converges in distribution and its limit coincides with the limit of the sequence $\{k!\binom{m}{k} n^{-k/2}\sigma_k^n \pi_{k\,m}^P h(X_1, \ldots, X_n)\}$.

This limit theorem is due to Rubin and Vitale (1980), but the case m=2 goes back to Hoeffding (1948a) (non-degenerate case) and von Mises (1947a) (degenerate case).

## 4. The bootstrap CLT for U-statistics.

We start by fixing the notation. Given $\{X_i\}_{i=1}^{\infty}$ i.i.d. with values in (S,S,P) let $P_n(\omega) = n^{-1}\sum_{i=1}^{n} \delta_{X_i(\omega)}$ be their empirical measure. Let $X_{n1}^{\omega} \ldots, X_{nn}^{\omega}$ be i.i.d. random variables with law $P_n(\omega)$. In the cited paper Bretagnolle studied the bootstrap of U-statistics. He proved that in the degenerated case the bootstrap works in probability if the bootstrap sample $m_n$ satisfies $m_n/n \to 0$ and a.s. if $m_n(\log n)^b/n \to 0$ for some $b > 1$, under quite strong moments conditions on h. He also observes that for $m_n = n$ this 'naive' bootstrap does not work for $h(x,y) = xy$ if $\mathbf{E}X_1 = 0$. This example gives a clue on both, why the naive bootstrap fails and how to proceed. Let $\mathbf{E}X_1 = 0$. Then

$$nU_2^n(h,\mathbf{P}) = n\binom{n}{2}^{-1}\sum_{i<j} X_iX_j = (n-1)^{-1}\left(\sum_{i=1}^{n} X_i\right)^2 - (n-1)^{-1}\left(\sum_{i=1}^{n} X_i^2\right). \tag{4.1}$$

By the CLT and LLN (4.1) converges in law to $\mathbf{E}X^2(g^2-1)$ where g in $N(0,1)$. But $nU_2^n(h,\mathbf{P}_n) = n\binom{n}{2}^{-1}\sum_{i<j} X_{n\,i}X_{n\,j} = (n-1)^{-1}(\sum_{i=1}^{n} X_{n\,i})^2 - (n-1)^{-1}(\sum_{i=1}^{n} X_{n\,i}^2)$ ,even after centering, is not a natural "replica" of $nU_2^n(h,P)$. In this case it is clearly more natural to apply the bootstrap CLT and the the bootstrap law of large numbers to the right side of (4.1) to obtain the bootstrap statistic that works is $(n-1)^{-1}(\sum_{i=1}^{n} X_{n,i} - \hat{\mathbf{E}}X_{n,i})^2 - (n-1)^{-1}(\sum_{i=1}^{n} X_{n,i} - \hat{\mathbf{E}}X_{n,i})^2)$.

This expression equals $nU_2^n(\bar{h}_n, P_n)$ where $\bar{h}_n(x,y) = h(x,y) - P_nh(x,\cdot) - P_nh(\cdot,y) + P_n^2h(\cdot,\cdot)$ $\bar{h}_n$ is the second term in the von Mises expansion of h with respect to $\mathbf{P}_n$. If h is of order k with respect to P let $\bar{h}_n$ be the k-th term in the von Mises expansion of h with repect to $P_n$. Then we take as the bootstrap statistics not $nU_2^n(h,P_n)$ but $nU_2^n(\bar{h}_n,P_n)$. Next we prove that this bootstrap procedure works a.s., for any bootstrap sample size $m_n \to \infty$ under minimal conditions.

**Lemma 4.1.** Let $\phi_j \in L_\infty(\mathrm{S},\mathcal{S}\ ,P)$ with $P\phi_j = 0$ for $j = i, \ldots, J$. Then $\omega$-a.s.

(4.2) $\lim_{n\to\infty}\hat{\mathcal{L}}\ (n^{-k/2}\sigma^n_k\pi^{P_n}_{k\,k}\ h^{\phi_j}_k(X_{n\,1},\ldots,X_{n\,n}) : j = 1,\ldots,J) =$

$\lim_{n\to\infty}\mathcal{L}\ (n^{-k/2}\sigma^n_k\qquad h^{\phi_j}_k(X_1\ \ ,\ldots,X_n\ \ ) : j = 1,\ldots,J).$

**Proof.** We have $\pi^{P_n}_{k\,k}\ h^{\phi_j}_k(x_1,\ldots,x_k) = (\phi_j(x_1) - P_n(\phi_j))\cdots(\phi_j(x_k) - P_n(\phi_j))$.

Hence by (3.1) $n^{-k/2}\sigma^n_k\pi^{P_n}_{k\,k}\ h^{\phi_j}_k(X_{n\,1},\ldots,X_{n\,n}) =$

$P_k(n^{-1/2}\sum_{i=1}^n(\phi_j(X_{n\,i})-P_n\phi_j), n^{-1}\sum_{i=1}^n(\phi_j(X_{n\,i})-P_n\phi_j)^2,\ldots,n^{-k/2}\sum_{i=1}^n(\phi_j(X_{n\,i})-P_n\phi_j)^k)$.

By the bootstrap CLT and the bootstrap law of large numbers in $\mathbf{R}^J$ (Bickel-Freedman (1981))

$\lim_{n\to\infty}\hat{\mathcal{L}}\ (n^{-1/2}\sum_{i=1}^n(\phi_j(X_{n\,i}) - P_n\phi_j) : j = 1,\ldots,J) =$

$\lim_{n\to\infty}\mathcal{L}\ (n^{-1/2}\sum_{i=1}^n \phi_j(X_i) : j = 1,\ldots,J) = \mathcal{L}\ (G_P(\phi_j) : j = 1,\ldots,J)$ a.s. ,

$P_n\phi_j \to P\phi_j$ a.s. and for $s \geq 2$ $\ n^{-1}\sum_{i=1}^n \phi^s_j(X_{n\,i}) \to P\phi^s_j$ in $\hat{\mathrm{p}}\mathrm{r}$, a.s.

Since polinomial commute with weak limits, we obtain that $\omega$-a.s.

$\lim_{n\to\infty}\hat{\mathcal{L}}\ (n^{-k/2}\sigma^n_k\pi^{P_n}_{k\,k}\ h^{\phi_j}_k(X_{n\,1},\ldots,X_{n\,n}) : j = 1,\ldots,J) =$

$\mathcal{L}\ (P_k((G_P(\phi_j),\sigma^2_{\phi_j},0,\ldots,0)) : j = 1,\ldots,J)$. $\square$

**Lemma 4.2.** Let f be a P-canonical function in $L_2(\mathbf{S}^k,\mathcal{S}^{\ k},P^k)$ satisfying the integrability condition: for each $\{i_1,\ldots,i_k\} \in N^k_{\underset{\sim}{k}}$

(4.3) $\mathbf{E}|f(X_{i_1},\ldots,X_{i_k})|^{2d/k} < \infty$ if $d = \#\{i_1,\ldots,i_k\}$

then a.s.

(4.4) $\lim_{n\to\infty}\hat{\mathcal{L}}\ (n^{-k/2}\sigma^n_k\pi^{P_n}_{k\,k}\ f(X_{n\,1},\ldots,X_{n\,n})) = \lim_{n\to\infty}\mathcal{L}\ (n^{-k/2}\sigma^n_k\ f(X_1,\ldots,X_n))$.

**Proof.** By (3.5) f is the limit in $L_2$ of functions of the form $g = \sum_{j=1}^J c_j h^{\phi_j}_k$. By the last lemma (4.4) holds for this functions. Now lemma 2.1-a) and b) with $P_n$ replacing $P$, gives

$\hat{\mathbf{E}}(n^{-k/2}\sigma^n_k\pi^{P_n}_{k\,k}(f-g))^2 = n^{-k}\binom{n}{k}\hat{\mathbf{E}}(\pi^{P_n}_{k\,k}(f-g))^2 \leq$

$k!^{-1}\hat{\mathbf{E}}(f-g)^2(X_{n\,1},\ldots,X_{n\,k}) = k!^{-1}n^{-k}\sum_{i_1,\ldots,i_k=1}^{n}(f-g)^2(X_{i_1},\ldots,X_{i_k})$ that converges a.s to $k!^{-1}\mathbf{E}(f-g)^2(X_{i_1},\ldots,X_{i_k})$ by (4.3). Now the lemma 2.2 gives the limit (4.4). □

This lemma is already a solution to the bootstrap problem if P is known: take $f=\pi_{k\,m}^{\mathbf{P}}h$ to obtain by theorem 3.1 and (4.7 $\lim_{n\to\infty}\hat{\mathcal{L}}\ (n^{-k/2}k!\binom{m}{k}\sigma_k^n\pi_{k\,k}^{P_n}\pi_{k\,m}^{P}\ \ h(X_{n\,1},\ldots,X_{n\,n})) =$
$\lim_{n\to\infty}\hat{\mathcal{L}}\ (n^{k/2}(U_m^n(h,P)-\mathbf{E}U_m^n(h,P))$

If $\mathbf{P}$ is not known we need the following lemma

**Lemma 4.3.** Let $k\le m$ and let $h(x_1,\ldots,x_m)$ be a symmetric measurable function satisfying the following integrability conditions:

(4.5) if $\#\{i_1,\ldots,i_k\}=p$ then $\mathbf{E}|\mathbf{P}^{m-k}h(X_{i_1},\ldots,X_{i_k},\ldots)|^{2p/k}<\infty$,

(4.6) if $\#\{i_1,\ldots,i_m\}=q$ then $\mathbf{E}|h(X_{i_1},\ldots,X_{i_m},\ldots)\mathbf{P}^{m-k}h(X_{i_1},\ldots,X_{i_k},\ldots)|^{q/m}<\infty$ and

(4.7) if $\#\{i_1,\ldots,i_{2m-k}\}=s$ then

$$\mathbf{E}|h(X_{i_1},\ldots,X_{i_m})h(X_{i_1},\ldots,X_{i_k},X_{i_{m+1}},\ldots,X_{i_{2m-k}})|^{s/2m-k}<\infty$$

then

(4.8) $\lim_{n\to\infty}\hat{\mathbf{E}}[n^{-k/2}\sigma_k^n(\pi_{k\,m}^{P_n}-\pi_{k\,k}^{P_n}\pi_{k\,m}^{P})\ \ h(X_{n\,1},\ldots,X_{n\,n})]^2 = 0\ \omega$–a.s.

**Proof.** By lemma 2.1 (a),( b) and (d) we have

(4.9) $\hat{\mathbf{E}}[n^{-k/2}\sigma_k^n(\pi_{k\,m}^{\mathbf{P}_n}-\pi_{k\,k}^{P_n}\pi_{k\,m}^{P})(h)(X_{n\,1},\ldots,X_{n\,n})]^2 =$

$n^{-k}\binom{n}{k}\hat{\mathbf{E}}[(\pi_{k\,m}^{P_n}-\pi_{k\,k}^{\mathbf{P}_n}\pi_{k\,m}^{\mathbf{P}})(h)(X_{n\,1},\ldots,X_{n\,n})]^2$

$\le \frac{1}{k!}\hat{\mathbf{E}}[\pi_{k\,m}^{P_n}(P_n^{m-k}-P^{m-k})h)(X_{n\,1},\ldots,X_{n\,k})]^2 =$

$\frac{1}{k!}\hat{\mathbf{E}}[(P_n^{m-k}h)^2(X_{n\,1},\ldots,X_{n\,n})-2(P_n^{m-k}h)(P^{m-k}h)(X_{n\,1},\ldots,X_{n\,n})+(\mathbf{P}^{m-k}h)^2(X_{n\,1},\ldots,X_{n\,n})] =$

$\frac{1}{k!}[\hat{\mathbf{E}}h(X_{n\,1},\ldots,X_{n\,k},\ldots,X_{n\,k+1},\ldots,X_{n\,m})h(X_{n\,1},\ldots,X_{n\,k},\ldots,X_{n\,m+1},\ldots,X_{n\,2m-k})-$

$2\hat{\mathbf{E}}h(X_{n\,1},\ldots,X_{n\,m})P^{m-k}h(X_{n\,1},\ldots,X_{n\,k},\cdots)+\hat{\mathbf{E}}(P^{m-k}h(X_{n\,1},\ldots,X_{n\,k},\cdots)^2]$.

The three summands in the last term of (4.9) are three V-statistics that satisfy the integrability condition in the law of large numbers (2.10). So the first and the third summands converge to $\mathbf{E}h(X_1,\ldots,X_m)h(X_1,\ldots,X_k,X_{m+1},\ldots,X_{2m-k})$ and the second converges to minus twice this quantity. Hence (48) is proved. □

It is obvious that the integrability condition (4.6) for $f=\pi^{\mathbf{P}}_{k,m}h$ is implied by condition (4.5) on h. Hence, the last three lemmas together with the CLT (3.1) yield:

**Theorem 4.4.** Let $(S,\mathcal{S},P)$ be a probability space and let $h:\mathbf{S}^m\to\mathbf{R}$ be a measurable symmetric function of canonical order k with respect to P. Let $\{X_i\}_{i=1}^{\infty}$ be i.i.d. r v. with values in $(S,\mathcal{S},P)$ and for each n let $X^{\omega}_{nj}=X_{nj}$, $j=1,\ldots,n$ be i.i.d. $(P_n(\omega))$, where $P_n(\omega)=n^{-1}\sum_{i=1}^{n}\delta_{X_i}(\omega)$. Let h satisfy the integrability conditions (4.8) to (4.10). Then the following sequences of probability distributions converge weakly and all have the same limits:

(a) $\{\mathcal{L}\,(n^{k/2}(U^n_m(h,P)-\mathbf{E}U^n_m(h,P)))\}_{n=1}^{\infty}$

(b) $\{\mathcal{L}\,(k!\binom{m}{k}n^{-k/2}\sigma^n_k\pi^P_{k\,m}h(X_1,\ldots,X_n))\}_{n=1}^{\infty}$

(c) $\{\hat{\mathcal{L}}\,(k!\binom{m}{k}n^{-k/2}\sigma^n_k\pi^{P_n}_{k\,k}\pi^P_{k\,m}(h(X_{n\,1},\ldots,X_{n\,n}))\}_{n=1}^{\infty}$ $\omega$-a.s. and

(d) $\{\hat{\mathcal{L}}\,(k!\binom{m}{k}n^{-k/2}\sigma^n_k\pi^{P_n}_{k\,m}(h(X_{n\,1},\ldots,X_{n\,n}))\}_{n=1}^{\infty}$ $\omega$-a.s.

**Remark 4.5.** (Another alternative) A slighty different (but essentially equivalent to the above) rationale for the bootstrap of degenerate U-statistics goes as follows: since $U^n_m(h,P)$ is of canonical order k and this is so crucial that it must be preserved when bootstrapping, the bootstrap statistics should be not $U^n_m(h,P_n)$ but instead $U^n_m(\hat{h}_n,P)$ where $\hat{h}_n$ is the k $P_n$ component of h namely

(4.9) $\hat{h}_n(\omega) = h - \sum_{i=0}^{k-1} \sigma_i^m \pi_{k\,m}^{P_n} h.$

Notice that by lemma 2.1 (e), for $0 \leq r < k$ $\pi_{r\,m}^{P_n} \hat{h}_n = \pi_{k\,m}^{P_n} h - \pi_{r\,m}^{P_n} h = 0$ and for $r \geq k$ $\pi_{r\,m}^{\mathbf{P}_n} \hat{h}_n = \pi_{r\,m}^{P_n} h$, so that

(4.10) $U_m^n(\hat{h}_n, P_n) = \sum_{r=k}^{n} \binom{m}{r} \binom{n}{r}^{-1} (\sigma_r^n \pi_{k\,m}^{P_n} h)(X_{n\,1}, \ldots, X_{n\,n}).$

The leading term in (4.13), multiplied by $n^{k/2}$ is equivalent to the term d) in theorem 4.5 and the remainig terms are $o(n^{k/2})$ by the the proof of theorem provided that the integrability conditions (4.5)-(4.7) hold not only for k, but for all $k' \leq k$. Hence, under this extra integrability, we have that the sequence $\{\mathcal{L}\ (n^{k/2}(U_m^n(\bar{h}_n, \mathbf{P}_n))\}_{n=1}^{\infty}$ has a.s. the same weak limit as the sequences a)-d) in theorem 4.5.

**Remark 4.6.** (Different bootstrap sample sizes). Since the CLT and the LLN in $\mathbf{R}^J$ can be bootstrapped for any bootstrap sample size $N_n \rightarrow \infty$ and since the $L_2(P_n)$ estimates of lemmas 4.2 4.3 also hold for bootstrap sample size $N_n \rightarrow \infty$. The sequences c) and d) are now, for $X_{n\,1}, \ldots, \mathbf{X}_{n\,N_n}$ i.i.d. $(\mathbf{P}_n)$, $\{\hat{\mathcal{L}}\ (k!\binom{m}{k} n^{-k/2} \sigma_k \pi_{k\,k}^{P}(h(X_{n\,1}, \ldots, X_{n\,N_n}))\}_{n=1}^{\infty}$ and $\{\hat{\mathcal{L}}\ (k!\binom{m}{k} n^{-k/2} \sigma_k^{N_n} \pi_{k\,k}^{P_n}(h(X_{n\,1}, \ldots, X_{n\,N_n}))\}_{n=1}^{\infty}$.

**Remark 4.7.** (Joint convergence) If $h_j \in L_2(\mathrm{S}^m, \mathcal{S}^{\ m}, P^m)$ for $j = i, \ldots, J$ and they have canonical orders $k_1, \ldots, k_J$ then the limit in (3.6) holds jointly. It also true that all this lemmas about bootstrap hold jointly. So U-statistics can be bootstrapped jointly.

## 5. The bootstrap of the law of large numbers for U-statistics.

The purpose of this section is to prove with as less moment condtions as possible that

(5.1) $U_m^n(h, P_n) \xrightarrow{P} \mathbf{E}h$ a.s.

**Theorem 5.1.** Suppose that for each possible combiation of integers $i_1, \dots, i_m$ $\mathbf{E}|h(X_{i_1}, \dots, X_{i_m})|^{\#\{i_1,\dots,i_m\}/m} < \infty$ then $U_m^n(h, P_n) \to \mathbf{E}h(X_1, \dots, X_m)$ a.s.

**Proof.** By (2.5) $U_m^n(h, P_n) = \sum_{j=0}^{m} \binom{m}{j} U_j^n(\pi_{j\,m}^{P_n} h, P_n)$.

The 0-th term is $P_n^m h$ and by the law of large numbers for V-statistics as in (2.10) $P_n^m \to P^m h$.

For $1 \leq j \leq m$ $U_j^m(\pi_{j\,m}^{P_n} h, P_n) = U_j^m(\pi_{j\,m}^{P_n} h\, I_{|h| \leq n^{1/2}}, P_n) + U_j^m(\pi_{j\,m}^{P_n} h\, I_{|h| > n^{1/2}}, P_n)$.

Now $\hat{V}ar(U_j^n(\pi_{j\,m}^{P_n} h\, I_{|h| \leq n^{1/2}}, P_n) = \binom{n}{j}^{-1} \hat{\mathbf{E}}((\delta_{X_{n\,1}} - P_n) \cdots (\delta_{X_{n\,j}} - P_n) P_n^{m-j} h\, I_{|h| \leq n^{1/2}})^2$

$\leq \binom{n}{j}^{-1} \hat{\mathbf{E}}(\delta_{X_{n\,1}} \cdots \delta_{X_{n\,j}} P_n^{m-j} h\, I_{|h| \leq n^{1/2}})^2 \cong$

$n^{-2m} \sum_{\alpha,\beta,\gamma=1}^{n} h(X_{\alpha_1}, \dots, X_{\alpha_j}, X_{\beta_1}, \dots, X_{\beta_{m-j}}) h(X_{\alpha_1}, \dots, X_{\alpha_j}, X_{\gamma_1}, \dots, X_{\gamma_{m-j}}) I_{|h| \leq n^{1/2}}$

$\cong n^{\frac{1}{2}-m-j} \sum_{\alpha,\beta=1}^{n} |h(X_{\alpha_1}, \dots X_{\alpha_j}, \dots, X_{\beta_1}, \dots, X_{\beta_{m-j}})|$

$= n^{1/2-m} \sum_{i_1 \dots, i_m}^{n} |h(X_{i_1}, \dots, X_{i_m})| \to 0$ a.s. by the law of large numbers for V-statistics.

As far as the second term $\hat{\mathbf{E}}|U_j^n(\pi_{j,m}^{P_n} h, P_n)| = \hat{\mathbf{E}}|(\delta_{X_{n\,1}} - P_n) \cdots (\delta_{X_j} - P_n) P_n^{m-j} h I_{|h| \geq n^{1/2}}|$.

Developing this last expression we need that $\hat{\mathbf{E}}|\delta_{X_{n\,1}} \cdots \delta_{X_{n\,j}} P_n^{m-j} h I_{|h| > n^{1/2}}| \to 0$ a.s.

But $\hat{\mathbf{E}}|\delta_{X_{n\,1}} \cdots \delta_{X_{n\,j}} P_n^{m-j} h I_{|h| > n^{1/2}}| \leq n^{-m} \sum_{i_1,\dots,i_m=1}^{n} h(X_{i_1}, \dots, X_{i_m})|\, I_{|h| > n^{1/2}} \to 0$ a.s. by the law of the large numbers and the argument in (2.3), Chapter I. □

Note that the result generalizes the bootstrap of the large numbers.

## 6. Multisample U-Statistics.

Suppose we have $c < \infty$ probability measures on **S** $P^{(1)}, \dots, P^{(c)}$ and c independent sample $X_1^{(1)}, \dots, X_{n_1}^{(1)}; \dots; X_1^{(c)}, \dots, X_{n_c}^{(c)}$ one from each $P^{(i)}$, with $n = n_1 + \cdots + n_c$ and $n_i/n \to \tau_i \in (0, \infty)$ for $1 \leq i \leq c$. Suppose we have also a function $h(x_1^{(1)}, \dots, x_{m_1}^{(1)}; \dots; x_1^{(c)}, \dots, x_{m_c}^{(c)})$ symmetric within each of its blocks of arguments.

For $n_1 \geq m_1, \dots, n_c \geq m_c$ the c-sample U-statistics corresponding to h and $P = (P^{(1)}, \dots, P^{(c)})$ is defined as

(6.1) $U^{n_1,\dots,n_c}_{m_1,\dots,m_c}(h,\mathbf{P}) = \prod_{j=1}^{c}\binom{n_j}{m_j}^{-1}\sum_{\alpha_1}\cdots\sum_{\alpha_c} h(X^{(1)}_{\alpha_1(1)},\dots,X^{(1)}_{\alpha_1(m_1)};\dots;X^{(c)}_{\alpha_c(1)},\dots,X^{(c)}_{\alpha_c(m_c)})$

where the sum runs over all the possibilities $1 \le \alpha_j(1) < \cdots < \alpha_j(m_j) \le n_j$ for $j = 1,\dots,c$.

Let $(\sigma^{n_1}_{m_1},\dots,\sigma^{n_c}_{m_c})h(x^{(1)}_1,\dots,x^{(1)}_{n_1};\dots;x^{(c)}_1,\dots,x^{(c)}_{n_c}) =$

$$\sum_{\alpha_1}\cdots\sum_{\alpha_c} h(x^{(1)}_{\alpha_1(1)},\dots,x^{(1)}_{\alpha_1(m_1)};\dots;x^{(c)}_{\alpha_c(1)},\dots,x^{(c)}_{\alpha_c(m_c)}).$$

Define $(\pi^{P^{(1)}}_{i_1,m_1},\cdots,\pi^{P^{(c)}}_{i_c,m_c})$ accordingly.

The Hoeffding decomposition holds in the 1-sample case,

(6.2) $U^{n_1,\dots,n_c}_{m_1,\dots,m_c}(h,P) =$

$$\sum_{i_1=0}^{m_1}\cdots\sum_{i_c=0}^{m_c}\prod_{j=1}^{c}\binom{m_j}{n_j}\binom{n_j}{i_j}^{-1}(\sigma^{n_1}_{i_1},\dots,\sigma^{n_c}_{i_c})(\pi^{P^{(1)}}_{i_1,m_1},\dots,\pi^{P^{(c)}}_{i_c,m_c})h(X^{(1)}_1,\dots,X^{(c)}_{n_c}).$$

h has canonical order k if $(\pi^{P^{(1)}}_{i_1,m_1},\dots,\pi^{P^{(c)}}_{i_c,m_c})h = 0$ for all $(i_1,\dots,i_c)$ such that $1 \le i_1 + \cdots + i_c < k$ and is the largest integer for which this holds. Then only formal changes are needed in the proof of the CLT 3.1 to obtain:

**Theorem 6.1.** If $h \in L_2(\mathbf{S}^m, \mathcal{S}^{\,m}, (P^{(1)})^{m_1}\cdots(\mathbf{P}^{(c)})^{m_c})$ where $m = m_1 + \cdots + m_c$ then

(6.3) $n^{k/2}[U^{n_1,\dots,n_c}_{m_1,\dots,m_c}(h,P) - \mathbf{E}U^{n_1,\dots,n_c}_{m_1,\dots,m_c}(h,\mathbf{P})]$

converges in distribution and its limit coincides with the limit of the sequence

(6.4) $n^{-k/2}\sum_{i_1,\dots,i_c\ i_1+\cdots+i_c=k}\prod_{j=1}^{c}\binom{m_j}{n_j}\binom{n_j}{i_j}^{-1}(\sigma^{n_1}_{i_1},\dots,\sigma^{n_c}_{i_c})(\pi^{P^{(1)}}_{i_1,m_1},\dots,\pi^{P^{(c)}}_{i_c,m_c})h(X^{(1)}_1,\dots,X^{(c)}_{n_c})$

This limit theorem can also be bootstrapped exactly in the same way as the bootstrap for 1-sample U-statistics, i.e. by replacing, in the last expression, $P^{(1)},\cdots,P^{(c)}$ by $P^{(1)}_n,\cdots,P^{(c)}_n$ and the c samples $X^{(1)}_1,\dots,X^{(i)}_{n_i}$ by the corresponding bootstrap samples $X^{(i)}_{n_i,1},\dots,X^{(i)}_{n_i,n_i}$. And in fact the bootstrap sample sizes may be taken to be $N_i \to \infty$ arbitrarily.

## 7. The CLT and the Bootstrap CLT of V-statistics.

Given a symmetric measurable function $h : \mathbf{R}^m \to \mathbf{R}$ and a probability measure P on $\mathbf{R}$, the V-statistic of order m based on h and P is defined as

(7.1) $V_m^n(h, P) = n^{-m} \sum_{i_1,\dots,i_m=1}^n h(X_{i_1}, \dots, X_{i_m}) = P_n^m h$

where $\{X_i\}$ are i.i.d. (P). Every symmetric statistic, and $V_m^n(h, P)$ is one, admits a Hoeffding expansion into U-statistics with P-canonical kernels. So the CLT and the bootstrap CLT for $V_m^n$ can be deduced from the results in the previous sections. This is done for a different bootstrap CLT in Bretagnolle (1983). Here, we will apply the same principles of previous sections but not the results themselves: it is somewhat easier to decompose a V-statistic into V-statistics with P-canonical kernels (instead of U-statistics ) and work with these. As in (2.5) we have

(7.2) $V_m^n(h, P) = P_n^m h = ((P_n - P) + P)^m h = \sum_{j=0}^m \binom{m}{j} (P_n - \mathbf{P})^j P^{m-j} h =$

$\sum_{j=0}^m \binom{m}{j} n^{-j} \sum_{i_1,\dots,i_j=1}^n (\delta_{X_{i_1}} - \mathbf{P}) \cdots (\delta_{X_{i_j}} - P) P^{m-j} h = \sum_{j=0}^m \binom{m}{j} V_j^n(\pi_{j,m}^P h, P)$ . Now we must introduce some extra notation in order to account for repetion of indices. Given a partition $\mathrm{Q} = (A_1, \dots, A_q)$ of $N_m = \{1, \dots, m\}$, where $q = \#Q$ and $A_i \neq \emptyset$, we let, for $h : \mathbf{R}^m \to \mathbf{R}$,

(7.3) $J_Q(h)(x_1, \dots, x_q) = h(x_{i_1}, \dots, x_{i_q})$ where $i_j = k$ if $j \in A_k, k = 1, \dots, q$.

With this notation we obviously have

(7.4) $V_m^n(h, \mathbf{P}) = n^{-m} \sum_Q \sigma_q^n (q! S_q J_Q h)(\mathbf{X}_1, \dots, \mathbf{X}_n)$.

where the summation runs over all partions Q of $N_m$. This is a decomposition of $V_m^n$ into U-statistics. Following Filippova (1961) we let

$\tilde{L}_2(\mathbf{R}^m, \mathcal{B}^m, \mathbf{P}^m) := \{h \in L_2(\mathbf{R}^m, \mathcal{B}^m, \mathbf{P}^m) : J_Q h \in L_2(\mathbf{R}^q, \mathcal{B}^q, \mathbf{P}^q)$ for each partition Q of $N_m\}$ and

(7.5) $||h||_{\tilde{L}^2} := (\sum_Q \mathbf{E}(J_Q h)^2(X_1, \dots, X_q))^{1/2}$.

**Theorem 7.1.** Let $H : \mathbf{R}^m \to \mathbf{R}$ be a symmetric measurable function. Then

(a) $P^n \sigma_m^n h = \binom{n}{m} P^m h$ for $m \leq n$.

(b) $V_k^n(\pi_{k,m}^P h, P) = (P_n - P)^k P^{m-k} h = P_n^k(\pi_{k,m}^P h), k \leq m.$

(c) If $h \in \tilde{L}_2(P^m)$ then $\mathbf{E}[n^{m/2}(P_n - P)^m h]^2 \leq c_m^2 ||h||_{\tilde{L}_2}^2$ where $c_m$ depends only on m.

**Proof.** (a) and (b) are trivial. (c) is proved in Filippova (Mises lemma). □

Our goal is to bootstrap the CLT for $V_m^n(h, P)$ for h with P-canonical order k. In this case the sum (7.2) becomes

(7.6) $V_m^n(h, P) - \mathbf{E}V_m^n(h, P) = \sum_{r=k}^m \binom{m}{r} V_r^n(\pi_{r,m}^P h, P).$

We show that the first term is the leading term:

**Lemma 7.2.** If $h \in \tilde{L}_2(P^m)$ has P-canonical order k, $1 \leq k \leq m$, then

(7.7) $n^k \mathbf{E}[V_m^n(h, \mathbf{P}) - \mathbf{E}V_m^n(h, P) - \binom{m}{k} V_k^n(\pi_{k,m}^P h, P)]^2 = O(n^{-1}).$

**Proof.** By lemma 7.1 (b), (7.6) and (7.6) $n^{k/2}||V_m^n(h, P) - \mathbf{E}V_m^n(h, P) - \binom{m}{k} V_k^n(\pi_{k,m}^P h, P)||_2 \leq \sum_{r=k+1}^m n^{k/2} \binom{m}{r} ||V_r^n(\pi_{r,m}^P h, P)||_2 \leq \sum_{r=k+1}^m n^{(k-r)/2} \binom{m}{r} c_r ||P^{m-k} h||_2.$ □

Filippova (1961), theorem 4, proves that if $h \in \tilde{L}_2(P^m)$ then

(7.8) $\mathcal{L}\,(n^{k/2}(P_n - P)^k h) \to_w \mathcal{L}\,(\,\int_0^1 \cdots \int_0^1 h(\mathbf{F}^{-1}(x_1), \ldots, \mathbf{F}^{-1}(x_k))\, d\beta(x_1) \cdots d\beta(x_k)\,)$

where $\beta$ is the Brownian bridge and $\mathbf{F}$ is the cumulative distribution function of $\mathbf{P}$. Call this law $\mu(h, P, k)$ then (7.7) and (7.8) give the CLT for V-statistics:

(7.9) If $h \in \tilde{L}_2(P^m)$ has P-canonical order k, $1 \leq k \leq m$, then

$\lim_{n\to\infty} \mathcal{L}\,(n^{k/2}(V_m^n(h, P) - \mathbf{E}V_m^n(h, P)) = \lim_{n\to\infty} \mathcal{L}\,(n^{k/2} \binom{m}{k} V_m^n(\pi_{k\,m}^P h, P)) = \mu(\binom{m}{k} \pi_{k,m}^P h, P, k).$

So, we have a situation analogous to that of section 3 and will bootstrap in the same way. Let, as before, $X^{\omega}_{n1}\dots,X^{\omega}_{nn}$ be i.i.d. random variables with law $P_n(\omega)$ and let $\hat{P}^{\omega}_n = n^{-1}\sum_{i=1}^{n}\delta_{X^{\omega}_{n1}}$ be the empirical measure of the bootstrap sample. We drop the variable $\omega$.

**Lemma 7.4.** Let $-\infty = t_0 < t_1 < \cdots < t_{d-1} < t_d = \infty$ and let $A_j = (t_{j-1}, t_j]$, $j = 1, \cdots, d$ be the associated partition of $\mathbf{R}$. For constants $g_{j_1,\cdots j_d}$ let $g(x_1,\cdots,x_d)$ be the function

(7.10) $g(x_1,\cdots,x_k) = \sum_{j_1,\cdots,j_k=1}^{d} g_{j_1,\cdots j_d} I_{A_{j_1}}(x_1)\cdots I_{A_{j_d}}(x_d)$

then

(7.11) $\lim_{n\to\infty}\hat{\mathcal{L}}\,(n^{k/2}(\hat{\mathbf{P}}_n - P_n)^k g) = \lim_{n\to\infty}\mathcal{L}\,(n^{k/2}(P_n - P)^k g)$ a.s.

**Proof.** By the bootstrap CLT in $\mathbf{R}^d$ (Bickel and Freedman (1981))

$\lim_{n\to\infty}\hat{\mathcal{L}}\,(n^{1/2}(\hat{P}_n - P_n)(A_i) : i = 1,\dots,d) = \lim_{n\to\infty}\mathcal{L}\,(n^{1/2}(P_n - P)(A_i) : i - 1,\dots,d))$ a.s.

So the result follows because

$n^{k/2}(P_n - P)^k g = \sum_{j_1,\cdots,j_k=1}^{d} g_{j_1,\cdots,j_k} n^{1/2}(P_n - P)(A_{j_1})\cdots n^{1/2}(P_n - P)(A_{j_d})$ and
$n^{k/2}(\hat{P}_n - P_n)^k g = \sum_{j_1,\cdots,j_k=1}^{d} g_{j_1,\cdots,j_k} n^{1/2}(\hat{P}_n - P_n)(A_{j_1})\cdots n^{1/2}(\hat{P}_n - P_n)(A_{j_k})$ and polinomials commute with weak limits. □

**Lemma 7.5.** Let $f \in \tilde{L}_2(\mathbf{P}^k)$ be a P-canonical function. Then $\omega$-a.s.

(7.12) $\lim_{n\to\infty}\hat{\mathcal{L}}\,(n^{k/2}(\hat{P}_n - P_n)^k g) = \lim_{n\to\infty}\mathcal{L}\,(n^{k/2}(P_n - P)^k g)$.

**Proof.** The set of functions g of the form (7.10) is dense in $\tilde{L}_2(P^k)$. Let then $f = \lim_{\tau\to 0} g_\tau$ in $\tilde{L}_2(\mathbf{P}^k)$ with $g_\tau$ of the form (7.10). By lemma 7.1(c)

(7.13) $\mathbf{E}(n^{k/2}(\hat{P}_n - P_n)^k(f - g_\tau))^2 \le c_k^2 \|f - g_\tau\|^2_{\tilde{L}_2(P^m)}$.

Moreover, again by lemma 6.1 (c)

(7.14) $\hat{\mathbf{E}}(n^{k/2}(\hat{P}_n - P_n)^k(f - g_\tau))^2 \le c_k^2 \|f - g_\tau\|^2_{\tilde{L}_2(P/\frac{m}{n})} = c_k^2 \sum_Q \|J_Q(f - g_\tau)\|^2_{\tilde{L}_2(P_n^q)}$

The last random variable in (6.14) is a sum of V-statistics that satisfy condition (4.3) ( actually, with some room left). (Note that symmetry is not a problem: collect terms in this sum corresponding to permutation-equivalent partitions). Hence, the law of large numbers (4.3) shows that

(7.15) $\limsup_{n\to\infty} \hat{\mathbf{E}}(n^{k/2}(\hat{P}_n - P_n)^k(f - g_\tau))^2 \leq c_k^2 \|f - g_\tau\|_{\tilde{L}_2}^2$ a.s.

Now the result follows from lemma 7.4 (7.13) and (7.15). □

Lemma 7.5 for $f = \pi_{k,m}^{\mathbf{P}} h$ and the CLT (3.9) give that if $h \in \tilde{L}_2(\mathbf{P}^m)$ then $\omega$ a.s.

(7.16) $\lim_{n\to\infty} \hat{\mathcal{L}}\ (n^{k/2}\binom{m}{k}(\hat{P}_n - P_n)^k \pi_{k,m}^P h) = \lim_{n\to\infty} \mathcal{L}\ (n^{k/2}(V_m^n(h,P) - \mathbf{E}V_m^n(h,P))$

Now we must replace $\pi_{k,m}^P$ by $\pi_{k,m}^{P_n}$. This is allowed by the following lemma:

**Lemma 7.6.** If $h(x_1,\cdots,x_m)$ is of order k with respect to k and $h \in \tilde{L}_2 P^m)$ then

(7.17) $\lim_{n\to\infty} \hat{\mathbf{E}}[n^{k/2}(P_n - P_n)^k(\pi_{k,m}^P - \pi_{k,m}^{P_n})h]^2 = 0$ $\omega$ a.s.

**Proof.** By lemma 2.1d and lemma 7.1c $\hat{\mathbf{E}}[n^{k/2}(\hat{P}_n - \mathbf{P}_n)^k(\pi_{k,m}^P - \pi_{k,m}^{P_n})h]^2 =$ $\hat{\mathbf{E}}n^{k/2}(\hat{P}_n - P_n)^k(P^{m-k} - P_n^{m-k})h]^2 \leq c_k^2 \|P^{m-k}h - P_n^{m-k}h\|_{\tilde{L}_2(P_n^k)}^2 =$ $c_k^2 \sum_Q \|J_Q(P^{m-k}h - P_n h)\|_{L_2(P_n^q)}^2$.

By developing the square as in the proof of lemma 2.4, and applying the law of large numbers for V-statistics (7.4) the lemma follows. □

The CLT (7.9) the limit (7.16) and the lemma 7.6 give the bootstrap CLT for V-statistics:

**Theorem 7.7.** Let $h(x_1,\cdots,x_m)$ be measurable symmetric function with P order k, $1 \leq k \leq m$, such that if $h \in \tilde{L}_2(P^m)$. Then the following sequences of laws converge weakly to the same limit:

(a) $\{\mathcal{L}\ (n^{k/2}(V_m^n(h,P) - \mathbf{E}V_m^n(h,\mathbf{P})\}_{n=1}^{\infty}$,

(b) $\{\mathcal{L}\ (n^{k/2}\binom{m}{k}V_k^n(\pi_{k,m}^P h, P)\}_{n=1}^{\infty}$,

(c) $\{\mathcal{L}\ (n^{k/2}\binom{m}{k}V_k^n(\pi_{k,k}^{P_n}\pi_{k,m}^P h, P_n)\}_{n=1}^{\infty}$ and

(d) $\{\mathcal{L}\ (n^{k/2}\binom{m}{k}V_k^n(\pi_{k,m}^{P_n} h, P_n)\}_{n=1}^{\infty}$.

**Remark 7.8.** As in the case of U-statistics, extensions of Theorem 6.7 are possible. In particular, the bootstrap sample size can be any $N_m \to \infty$. A particular V-statistics of interest is the k-th term in the Taylor expasion of a von Mises statistical functional, namely $(P_n - P)^m f = V_m^n(\pi_{m,m}^P f)$, so that this is the special case of (7.6) for k=m. Hence, the equivalence beetween (a) and (b) in Theorem 7.7 gives:

**Corollary 7.9.** (The bootstrap of Filippova's CLT). Let $f \in \tilde{L}_2(P^k)$ be a symmetric function then

(7.18) $\lim_{n\to\infty}\hat{\mathcal{L}}\ (n^{m/2}(\hat{\mathbf{P}}_n - P_n)^m f) = \lim_{n\to\infty}\mathcal{L}\ (n^{m/2}(P_n - P)^m f) = \mu(f, P, m)$ a.s.

## 8. Applications.

In this applications we will need the following lemma.

**Lemma 8.1.** If Z has a continuous law and $\hat{c}_n(\alpha) = \inf\{t : \hat{P}\{\hat{T}_n \geq t\} \geq \alpha\}$ and $\hat{T}_n \to_w Z$ in pr. and $T_n \to_w Z$ then $P\{T_n \geq \hat{c}_n(\alpha)\} \to \alpha$.

We omit the proof because it is, except for a change in notation, theorem 1 in Beran (1984).

**Example 8.2.** (Hoeffding test for independence) Hoeffding (1948b) proposed a test of independence based on

(8.1) $\Delta(\mathbf{F}) = \int_{-\infty}^{\infty} (\mathbf{F}(x,y) - \mathbf{F}(x,\infty)\mathbf{F}(\infty,y))^2 \, d\mathbf{F}(x,y)$

$\mathbf{F}$ is c.d.f of a r.v. $(X,Y)$ so $F(x,y) = P\{X \leq x \, Y \leq y\}$ and $F(x,\infty) = \mathbf{P}\{X \leq x\}$

Consider $H = \{P \in \mathcal{P} \ (\mathbf{R}^2) : P$ has a continuous joint and marginal densities $\}$

$H_0 = \{P \in H :$ P is the product of its densities densities $\}$

We want to test whether P is in h or in $H_0$.

Note that if $P \in H$ then $P \in H_0$ iff $\Delta(\mathbf{F}) = 0$. Note also that $\Delta(\mathbf{F}) = P^5 h$ where

$h((x_1,y_1),(x_2,y_2),(x_3,y_3),(x_4,y_4),(x_5,y_5)) = 4^{-1}\varphi(x_1,x_2,x_3)\varphi(x_1,x_4,x_5)\varphi(y_1,y_2,y_3)\varphi(y_1,y_4,y_5)$

and $\varphi(x_1,x_2,x_3) = I_{x_2 \leq x_1} - I_{x_3 \leq x_1}$.

So we can hope to get a test of independence from

(8.2) $\int_{-\infty}^{\infty} (\mathbf{F}_n(x,y) - \mathbf{F}_n(x,\infty)\mathbf{F}_n(\infty,y))^2 \, d\mathbf{F}_n(x,y) = V_5^n(h,P)$.

Since U and V-statistics are asymptotically equivalent and the convergence of U-statistics requires less moments we will use the U-statistics over the kernel h. Now if $P \in H_0$ then h has $\mathbf{P}$-canonical order 2 so $nU_5^n(h,P) \to \mathcal{L}\ (\theta)$ and $\binom{5}{2} nU_2^n(\pi_{2,5}^{P_n} h, P_n) \to \mathcal{L}\ (\theta)$.

Since $\theta$ is the limit of U statistics of order 2 there are constants $\lambda_j$ and a sequence of i.i.d. r.v. N(0,1) $g_j$ so that $\mathcal{L}\ (\theta) = \mathcal{L}\ (\sum_{j=1}^{\infty} \lambda_j (g_j^2 - 1))$ where $\sum_{j=1} \lambda_j^2 < \infty$. Hence has a density and by lemma 7.1 taking $\hat{c}_n(\alpha)$ $=\inf\{t : \hat{P}\{\binom{5}{2} nU_2^n(\pi_{2,5}^{P_n} h, P_n) \geq t\} > \alpha\}$ then $P\{nU_5^n(h,\mathbf{P}) \geq \hat{c}_n(\alpha)\} \to \alpha$ if $P \in H_0$. If $P \in H - H_0$ by the law of large numbers for U-statistics $U_5^n(h,P) \to \Delta(\mathbf{F}) \neq 0$. Hence $P\{nU^n h \geq \hat{c}_n(\alpha)\} \to 1$.

**Example 8.3.** ( A bootstrap test for symmetry about zero.) Take $H = \{P \in \mathcal{P} \ (\mathbf{R}) : \int x^2 \, dP < \infty\}$, and $H_0 = \{\mathbf{P} \in \mathcal{P} \ (\mathbf{R}) : P$ is symmetric about zero $\}$. Consider the statistic

(8.3) $T_n = \int_{(-\infty}^{\infty}(P_n(I_{(-\infty,x]} - I_{[x,\infty)}))^2\,dx = P_n^2 h$ where $h(u,v) = (|u| \wedge |v|)(I_{u,v>0} + I_{u,v<0} - I_{u>0,v<0 \text{ or } u<0,v>0})$ (a modifiction of an example in Filippova).

Now it is very easy to chek that the expansion of $T_n$ is

(8.4) $\int_{-\infty}^{\infty}[P(I_{(-\infty,x]} - I_{[x,\infty)})]^2 + 2[(P_n - P)(I_{(-\infty,x]} - I_{[x,\infty)})][P(I_{(-\infty,x]} - I_{[x,\infty)})] +$
$[(P_n - P)(I_{(-\infty,x]} - I_{[x,\infty)})]^2\,dx$

In the case that $\mathbf{P} \in H_0$

(8.5) $T_n = \int_{-\infty}^{\infty}[(P_n - P)(I_{-\infty,x]} - I_{[x,\infty)})]^2\,dx = (P_n - P)^2 h.$

The limiting distribution of $n(P_n - P)^2 h$ is the law of a shift of an infinity linear combination of centered independent chi-square random variables of order one, hence it is continuous. So, if we define $\hat{c}_n(\alpha)$ by $\inf\,\{t : \hat{\mathbf{Pr}}\{(\hat{P}_n - P_n)^2 h| > t\} \geq \alpha\}$ we have $\hat{\mathbf{Pr}}\{nT_n > \hat{c}_n(\alpha)\} \to \alpha$ if $P \in H_0$.

In the case that $P \in H - H_0$ $T_n \to \mathbf{P}^2 h$ that is diferent from zero by right continuity of $\mathbf{F}$. So $\hat{\mathbf{Pr}}\{nT_n > \hat{c}_n(\alpha)\} \to 1$ if $P \in H - H_0$.

These tests are also consistent against local alternatives of the form $P + \Delta n^{-\lambda}$ $P \in H_0$ and $P + \Delta \in H - H_0$.

# CHAPTER IV

# SOME BOOSTRAPS TESTS OF SYMMETRY FOR UNIVARIATE CONTINUOUS DISTRIBUTIONS

## 1. Introduction.

The following is an extended version of the paper Arcones-Giné (1989b) about tests of symmetry based on the bootstrap.

There are statistical procedures which are sensible to departures from symmetry, therefore requiring testing for symmetry in advance. In this chapter we propose some natural tests for symmetry whose initial values can (in fact, only) be computed using a modified bootstrap procedure. Given a probability P in $\mathbf{R}$, we say that P is symmetric if there exists some $\theta \in \mathbf{R}$ such that $\mathbf{P}(A) = \mathbf{P}(2\theta - A)$ for all Borelian sets A.

Given a sample $X_1, \ldots, X_n$ of the c.d.f. (cumulative distribution function) $\mathbf{F}$, we would like to decide whether $\mathbf{F}$ is symmetric or not. A good measure of how symmetric is $\mathbf{F}_n$ (and therefore F) is

(1.1) $\inf\{||\mathbf{F}_n - \mathbf{G}||_\infty : \mathbf{G}$ is a symmetric c.d.f.$\}$

where $|| \cdot ||_\infty$ denotes the Kolmogorov distance.

(1.2) $||\mathbf{F} - \mathbf{G}||_\infty = \sup_t |\mathbf{F}(t) - \mathbf{G}(t)|$ for c.d.f. $\mathbf{F}$ and $\mathbf{G}$.

It is known (see Schuster(1987) and Schuster-Narvarte (1973) ) that given a c.d.f. $\mathbf{F}$, there is a symmetric c.d.f. s$\mathbf{F}$ such that

(1.3) $||\mathbf{F} - s\mathbf{F}||_\infty = \inf\{||\mathbf{F} - \mathbf{G}||_\infty : \mathbf{G}$ is a symmetric c.d.f.$\}$.

Moreover s$\mathbf{F}$ is the symmetrized of F with respect to some parameter $\theta$.

This suggests to test for symmetry by rejecting the hypothesis that F is symmetric if $||\mathbf{F}_n - s\mathbf{F}_n||$ is too large. We know that $\sqrt{n}||\mathbf{F}_n - s\mathbf{F}_n||_\infty$ converges weakly. So we think of taking the quantiles of the limit as the initial values in our test. The problem is that the limit depends on the distribution.

The way this problem has been solved is by bootstrapping. Call $\mathbf{F}_n$ the bootstrap empirical distribution. We will get:

(1.4) $\sqrt{n}\|\mathbf{F}_n - s\mathbf{F}_n\|_\infty \to \mathbf{Z}_\mathbf{F}$ if $\mathbf{F}$ is symmetric, where $\mathbf{Z}_\mathbf{F}$ is a law depending on $\mathbf{F}$.

(1.5) $\sqrt{n}\|\mathbf{F}_n - s\mathbf{F}_n\|_\infty \to \infty$ pr. if $\mathbf{F}$ is not symmetric.

(1.6) $\sqrt{n}\|\mathbf{F}_{n,s\mathbf{F}_n} - s\mathbf{F}_{n,s\mathbf{F}_n}\|_\infty \to \mathbf{Z}_{s\mathbf{F}}$ for every $\mathbf{F}$.

So if we take $t_{n,\alpha}(\omega)$ =inf$\{t : \hat{\mathbf{Pr}}\ \{\sqrt{n}\|\mathbf{F}_{n,s\mathbf{F}_n} - s\mathbf{F}_{n,s\mathbf{F}_n}\| \geq t\} \geq \alpha\}$

$\mathbf{Pr}\{\sqrt{n}\|\mathbf{F}_n - s\mathbf{F}_n\|_\infty \geq t_{n,\alpha}\} \to \alpha$ if $\mathbf{F}$ is not symmetric.

$\mathbf{Pr}\{\sqrt{n}\|\mathbf{F}_n - s\mathbf{F}_n\|_\infty \geq t_{n,\alpha}\} \to 1$ if $\mathbf{F}$ is symmetric.

This is the idea in Romano (1988) about distance tests but his smothness conditions are too strong to apply to the present situation. In Schuster-Barker (1987) there is simulation of this tests but no analitical justification. In this chapter we give an analitical justification for this kind of test of symmetry and we give another 3 different alternatives.

## 2. A symmetric bootstrap central limit theorem.

Let $\mathcal{P}(\mathbf{R})$ be the set of probability measures on $\mathbf{R}$. A *parameter of location of symmetry* $\theta : \mathcal{P}(\mathbf{R}) \to \mathbf{R}$ is a funtion satisfying:

(1) if $\mathbf{P}$ is symmetric then $\theta$ $(\mathbf{P})$ is its center of symmetry and

(2) the function $\theta_n : \mathbf{R}^n \to \mathbf{R}$ defined by $\theta_n(x_1, \ldots, x_n) = \theta(n^{-1}\sum_{i=1}^n \delta_{x_i})$ is measurable.

Given a location parameter $\theta$ we define $s^\theta\mathbf{P}$ (sP if no confusion may arise) by

(2.1) $s^\theta\mathbf{P}(A) = \frac{1}{2}(\mathbf{P}(A) + \mathbf{P}(2\theta - A))$; $A \in \mathcal{B}$,

and $s^\theta\mathbf{F}$,$s^\theta f$, denote respectively the c.d.f and the density (if it exists) of $s^\theta\mathbf{P}$.Note that $s^\theta\mathbf{P}$ is the symmetrized probability measure of $\mathbf{P}$ about $\theta$. If no confusion is possible we will use $\theta$ for

$\theta(\mathbf{P})$, $\theta_n$ for $\theta(\mathbf{P}_n)$ and $\hat{\theta_n}$ for $\theta(\mathbf{P}_{n,\mathbf{P}_n})$, where $\mathbf{P}_{n,\mathbf{P}_n} = \mathbf{P}_{n,\mathbf{P}_n(\omega)}$ is the empirical measure of n i.i.d. random variables $\mathbf{Y}^{\omega}_{n1}, \cdots, \mathbf{Y}^{\omega}_{nn}$ with common probability law $s\mathbf{P}_n(\omega)$. The variable $\omega$ will often be omitted. Finally $\hat{\mathbf{Pr}}$,$\hat{\mathbf{E}}$ will denoted **Pr**, **E** conditionally on $\mathbf{P}_n(\omega)$.

We will use repeateadly the following lemma.

**Lemma 2.1.** Let $\mathcal{C} = \{(-\infty, t], (-\infty, t) : t \in \mathbf{R}\}$. Let $\mu_n$, $\mu$ be probabilities measures in **R** . If $\sup_t |(\mu_n - \mu)(-\infty, t]| \to 0$ then

(2.2) $n^{1/2}(\mathbf{P}_{n,\mu_n} - \mu_n)(\mathcal{C}) \to_w U(\mu(\mathcal{C}))$ in $l^\infty(\mathcal{C})$

**Proof.** We will deduce the result from corollary 2.7 in Giné-Zinn (1990a). Note first that the class $\mathcal{C}$ is uniformly pregaussian because it is a V-C class. Call $\mathcal{J} = \{J;\ J$ is an interval in $\mathbf{R}\}$. From that result what we must show is that $\sup_{J \in \mathcal{J}} |(\mu_n - \mu)J| \to 0$. Note that $\sup_{J \in \mathcal{J}} |(\mu_n - \mu)J| \leq 2 \sup_t |(\mu_n - \mu)(-\infty, t]) \to 0$. □

**Definition 2.2.** Let $\theta = \theta(\mathbf{P}), \mathbf{P} \in \mathcal{P}(\mathbf{R})$, be a parameter of location of symmetry and let $\Pi$ be class of probability measures such that if $\mathbf{P} \in \Pi$, also $s^\theta \mathbf{P} \in \Pi$. We say that $\theta$ is a strong bootstrap consistent parameter of location of symmetry for the class $\Pi$ if the following properties hold (with $X_i$ i.i..d. (**P**) ):

(i) $\lim_{n\to\infty} \theta_n(X_1, \ldots, X_n) = \theta(\mathbf{P})$ a.s. for all $\mathbf{P} \in \Pi$.

(ii) For all *symmetric* $\mathbf{P} \in \Pi$ there is a random variable $\tilde{\theta}(\mathbf{P})$ such that the following limits holds jointly weakly:

(2.3) $n^{1/2}(\mathbf{P}_n - \mathbf{P})(I_{(-\infty,t]}) \xrightarrow{\mathcal{L}} U(\mathbf{F}(t))$ in $l_\infty(\mathbf{R})$ where U is a Brownian bridge and

$$n^{1/2}(\theta(\mathbf{P}_n) - \theta(\mathbf{P})) \xrightarrow{\mathcal{L}} \tilde{\theta}(\mathbf{P})$$

(iii) (First note that lemma 2.1, (i) in this definition and the Glivenko-Cantelli theorem imply that $n^{1/2}(\mathbf{P}_{n,s\mathbf{P}_n(\omega)} - s\mathbf{P}_n(\omega))(I_{(-\infty,t]}) \xrightarrow{\mathcal{L}} U(s\mathbf{F}(t))$ in $l_\infty(\mathbf{R})$ a.s.)

For all $\mathbf{P} \in \Pi$ there is a variable $\tilde{\theta}(s\mathbf{P})$ such that the following limit holds jointly

(2.4) $n^{1/2}(\mathbf{P}_{n,s\mathbf{P}_n(\omega)} - s\mathbf{P}_n(\omega))(I_{(-\infty,t]}) \rightarrow_w U(s\mathbf{F}(t))$ in $l_\infty(\mathbf{R})$ and

$n^{1/2}(\theta(\mathbf{P}_{n,s\mathbf{P}_n(\omega)}) - \theta(\mathbf{P}_n(\omega)) \xrightarrow{\mathcal{L}} \tilde{\theta}(s\mathbf{P})$ a.s.

where sP=$s^\theta\mathbf{P}$ with $\theta = \theta(\mathbf{P})$ and s$\mathbf{F}_n = s^{\theta_n}\mathbf{F}_n$ with $\theta_n = \theta_n(X_1, \ldots, X_n)$.

Analogously we say that $\theta$ is a weak bootstrap coinsistent parameter of location if the limits in (2.4) are in pr.

Obviously, conditions (i)-(iii) in Definition 2.1 hold if $\theta = \theta(\mathbf{P})$ is sufficiently differentiable. At least in the case we are interested in, it seems more expeditive to check these conditions directly.

**Theorem 2.3.** Let $\Pi$ be a set of probability measures $\mathbf{P}$ on $\mathbf{R}$ which are absolutely continuous and whose density f are uniformly continuous on $\mathbf{D_P} := \{x : 0 < \mathbf{F}(x) < 1\}$. Let $\theta$ be a weak bootstrap consistent parameter of location of symmetry of the class $\Pi$. Then the following limits hold:

(i) For all symmetric $\mathbf{P} \in \Pi$

$\lim_{n\to\infty} n^{1/2}[\mathbf{F}_n - s\mathbf{F}_n](\omega)(.) = 2^{-1}\mathbf{U}(s\mathbf{F}(\cdot)) + 2^{-1}\mathbf{U}(1 - s\mathbf{F}(\cdot)) + \tilde{\theta}(\mathbf{P})(f)(\cdot)$ weakly in $l^\infty(D_{s\mathbf{P}})$.

(ii) For all $\mathbf{P} \in \Pi$,

$\lim_{n\to\infty} n^{1/2}[\mathbf{F}_{n,s\mathbf{P}_n(\omega)} - s\mathbf{F}_{n,s\mathbf{P}_n(\omega)}](\cdot) = 2^{-1}\mathbf{U}(s\mathbf{F}(.)) + 2^{-1}\mathbf{U}(1 - s\mathbf{F}(\cdot)) + \tilde{\theta}(s\mathbf{P})(sf)(\cdot)$ weakly in $l^\infty(D_{s\mathbf{P}})$ in pr.

(iii) Moreover if $\theta$ is strong bootstrap consistent parameter for all $\mathbf{P} \in \Pi$, such that the Hölder type condition holds

(2.5) $h(\delta) := \sup_{|t-s|\leq\delta}|\mathbf{F}(t) - \mathbf{F}(s)| \ln\ \ln(1/|t-s|) \to 0$ as $\delta \to 0$ then

$\lim_{n\to\infty} n^{1/2}[\mathbf{F}_{n,s\mathbf{P}_n(\omega)} - s\mathbf{F}_{n,s\mathbf{P}_n(\omega)}](\cdot) = Z := 2^{-1}\mathbf{U}(s\mathbf{F}(\cdot)) + 2^{-1}\mathbf{U}(1 - s\mathbf{F}(\cdot)) + \tilde{\theta}(s\mathbf{P})(sf)(\cdot)$

weakly in $1^\infty(D_{s\mathbf{P}})$ a.s.

**Proof.** We have that

(2.6) $n^{1/2}(\mathbf{F}_n - s\mathbf{F}_n)(t) = n^{1/2}\mathbf{P}_n(2^{-1}I_{(-\infty,t]} - 2^{-1}I_{[2\theta_n - t,\infty)}) =$

$n^{1/2}(\mathbf{P}_n - \mathbf{P})(2^{-1}I_{(-\infty,t]} - 2^{-1}I_{[2\theta_n - t,\infty)}) + n^{1/2}\mathbf{P}(2^{-1}I_{(-\infty,t]} - 2^{-1}I_{[2\theta_n - t,\infty)}) = I + II.$

Now $I = n^{1/2}(\mathbf{P}_n - \mathbf{P})(2^{-1}I_{(-\infty,t]} + 2^{-1}I_{(-\infty,2\theta_n - t)}) =$

$n^{1/2}(\mathbf{P}_n - \mathbf{P})(2^{-1}I_{(-\infty,t]} + 2^{-1}I_{(-\infty,2\theta - t]}) + n^{1/2}(\mathbf{P}_n - \mathbf{P})(2^{-1}I_{(-\infty,2\theta_n - t)} - 2^{-1}I_{(-\infty,2\theta - t]}) =$

$I_a + I_b$.

Since $\mathcal{C}$ is a V-C class, there is gaussian process $G_P$ such that $n^{1/2}(\mathbf{P}_n - \mathbf{P}) \to_w G_P$ in $1_\infty(\mathcal{C})$.

So for each $\epsilon > 0$

$\lim_{\delta\to 0} \limsup_{n\to\infty} \mathbf{Pr}\{ \sup_{\rho(C,D)\leq\delta}|n^{1/2}(\mathbf{P}_n - \mathbf{P})(I_C - I_D| > \epsilon\} = 0$

where $\rho^2(C,D) = Var(I_C(\mathbf{X}) - (I_D(\mathbf{X}))$.

Note that $\rho^2(I_{(-\infty,u]}, I_{(-\infty,v)}) = Var(I_{(-\infty,u]}(\mathbf{X}) - (I_{(-\infty,v)}(\mathbf{X})) \leq |F(u) - F(v)|$.

Since $\mathbf{F}$ is uniformly continuous for each $\epsilon > 0$

$\lim_{\delta\to 0} \limsup_{n\to\infty} \mathbf{Pr}( \sup_{|\theta-\theta'|\leq\delta}|n^{1/2}(\mathbf{P}_n - \mathbf{P})(I_{-\infty,\theta]} - I_{(-\infty,\theta')})| > \epsilon) = 0$

Since $\theta_n \to \theta$ a.s.

(2.7) $\sup_t|n^{1/2}(\mathbf{P}_n - \mathbf{P})(2^{-1}I_{(-\infty,2\theta_n - t)} - 2^{-1}I_{(-\infty,2\theta - t]})| \to 0$ in pr.

Since f is uniformly continuos and $\theta_n \to \theta$ a.s.

(2.8) $\sup_t|n^{1/2}\mathbf{P}(2^{-1}I_{(-\infty,t]} - 2^{-1}I_{(-\infty,2\theta - 2\theta_n - t]}) - n^{1/2}(\theta_n - \theta)f(t)| \to 0.$ a.s. Hence (2.6)-(2.8) give that the distribution of $n^{1/2}(\mathbf{F}_n - s\mathbf{F}_n)$ in $1^\infty(\mathbf{D_P})$ is asimptotically equivalent to the distribution of

(2.9) $n^{1/2}(\mathbf{P}_n - \mathbf{P})(2^{-1}I_{(-\infty,t]} + 2^{-1}I_{(-\infty,2\theta - t]}) + n^{1/2}f(t)(\theta_n - \theta)$

which converges weakly to $2^{-1}\mathbf{U}(\mathbf{F}(t))+2^{-1}\mathbf{U}(\mathbf{F}(1-t))+\tilde{\theta}f(t)$.

The proof of (ii) is similar

(2.10) $n^{1/2}[\mathbf{F}_{n,s\mathbf{P}_n}-s\mathbf{F}_{n,s\mathbf{P}_n}](I_{(-\infty,t]})=n^{1/2}(\mathbf{F}_{n,s\mathbf{P}_n})(2^{-1}I_{(-\infty,t]}-2^{-1}I_{[2\tilde{\theta}_n-t,\infty)})=$ $n^{1/2}[\mathbf{F}_{n,s\mathbf{P}_n}-s\mathbf{P}_n](2^{-1}I_{(-\infty,t]}-2^{-1}I_{[2\tilde{\theta}_n-t,\infty)})+n^{1/2}s\mathbf{P}_n(2^{-1}I_{(-\infty,t]}-2^{-1}I_{[2\tilde{\theta}_n-t,\infty)})=I'+II'$

Now I'=$n^{1/2}[\mathbf{F}_{n,s\mathbf{P}_n}-s\mathbf{P}_n](2^{-1}I_{(-\infty,t]}-2^{-1}I_{(-\infty,2\theta-t]}+$ $n^{1/2}[\mathbf{F}_{n,s\mathbf{P}_n}-s\mathbf{F}_{n,s\mathbf{P}_n}](2^{-1}I_{(-\infty,2\theta-t]}-2^{-1}I_{(-\infty,2\tilde{\theta}_n-t)})=I'_a+I'_b$. Similarly to (2.6) since $\tilde{\theta}\xrightarrow{\mathcal{L}}\theta$ a.s. $\sup_t|I'_b|\xrightarrow{P}0$ in pr.

As far as II'

(2.11) $II'=n^{1/2}\mathbf{P}_n(4^{-1}I_{(-\infty,t]}+4^{-1}I_{[2\theta_n-t,\infty)}-4^{-1}I_{(-\infty,2\theta_n-2\tilde{\theta}_n+t]}-4^{-1}I_{[2\theta_n-\bar{t},\infty)})=$ $=n^{1/2}[\mathbf{P}_n-\mathbf{P}](4^{-1}I_{(-\infty,t]}+4^{-1}I_{[2\theta_n-t,\infty)}-4^{-1}I_{(-\infty,2\theta_n-2\tilde{\theta}_n+t]}-4^{-1}I_{[2\tilde{\theta}_n-t,\infty)})+$ $n^{1/2}\mathbf{P}(4^{-1}I_{(-\infty,t]}+4^{-1}I_{[2\theta_n-t,\infty)}-4^{-1}I_{(-\infty,2\theta_n-2\tilde{\theta}_n+t]}-4^{-1}I_{[2\tilde{\theta}_n-t,\infty)})=II'_a+II'_b.$

Now for any $a_n$ such that $a_n\to 0$

(2.12) $\omega(a_n):=\sup_{|t-s|\leq a_n}|n^{1/2}(\mathbf{P}_n-\mathbf{P})(I_{(-\infty,t]}-I_{(-\infty,s]}|\to 0$ in pr.

Hence taking $a_n$, with $a_n\to 0$ and $\sqrt{n}a_n\to\infty$. For each $\epsilon>0$

(2.13) $\mathbf{P}\times\hat{\mathbf{P}}\{sup_t|II'_a(t)|>\epsilon\}\leq\mathbf{P}\{\omega(a_n)>\epsilon\}+\mathbf{P}\times\hat{\mathbf{P}}\{2|\theta_n-\hat{\theta}_n|>a_n\}\to 0.$

As for $II'_b$ using a.s. convergence of $\theta_n$ to $\theta$, the $\omega$ pr. stochastic boundedness of $n^{1/2}(\tilde{\theta}_n-\theta_n)$ and the uniformly diffferentiability of f, we have that $\omega$ pr.

(2.14) $||\frac{n^{1/2}}{4}\mathbf{P}(I_{(-\infty,t]}+I_{[2\theta_n-t,\infty)}-I_{(-\infty,2\theta_n-2\tilde{\theta}_n+t]}-I_{[2\tilde{\theta}_n-t,\infty)})-n^{1/2}(\tilde{\theta}_n-\theta_n)(sf(t)||_\infty\to$ $0$ *a.s.*

Collecting (2.10)-(2.14), (ii) follows.

As far as (iii), the proof is similar. Note that for all the remainders except $II_b$ now we have converges a.s. The problem is that (2.12) does not hold for any sequence $c_n\to 0$ That is why we ask the condition (2.5).

We need that for some $c_n\to 0$ with $\sqrt{n}c_n\to\infty$

(2.15) $\omega(c_n) = \sup_{|t-s|\leq c_n} |n^{1/2}(\mathbf{P}_n - \mathbf{P})(I_{(-\infty,t]} - I_{(-\infty,s]}| \to 0$ a.s.

Call $\xi_i = \mathbf{F}(\mathbf{X}_i)$ $\mathrm{U}_n = \sqrt{n}(n^{-1}\sum_{i=1}^{n}(I_{\xi_i \leq t} - t)\ )$ Then a.s. $U_n(\mathbf{F}(t)) = n^{1/2}(\mathbf{P}_n - \mathbf{P})(t)$.

By the Hölder condition in (2.5), for any $\lambda > 0$ if $|t-s| \leq n^{-\lambda}$ then

$|\mathbf{F}(t) - \mathbf{F}(s)| \leq h(n^{-\lambda})/\ln\ \ln n^{-\lambda} \leq h(n^{-\lambda})/\ln\ \ln n$. Hence

$\sup_{|t-s|\leq n^{-\lambda}} |n^{1/2}(\mathbf{P}_n - \mathbf{P})(s,t]| \leq \sup_{|u-v|\leq h(n^{-\lambda})/\ln\ \ln n} |U_n(u) - U_n(v)|$ .

By theorem 4 in Mason-Shorack-Wellner (1983)

$\limsup_{n\to\infty} \sup_{|u-v|\leq c/\ln\ln n} |U_n(u) - U_n(v)| \to \sqrt{2c}$ a.s. Hence for any $b_n \to 0$

(2.16) $\sup_{|u-v|\leq b_n/\ln\ln n} |U_n(u) - U_n(v)| \to 0$ a.s.

Therefore for any $\lambda > 0$

(2.17) $\sup_{|t-s|\leq n^{-\lambda}} |n^{1/2}(\mathbf{P}_n - \mathbf{P})(s,t]| \to 0$ a.s. □

**Remark 2.4.** Schuster and Barker (1987) suggest considering a smothed symmetric bootstrap. This requires proving that the limit in (i) of theorem 2.3 holds also for the statistic

(2.18) $n^{1/2}[\mathbf{P}_{n,s\mathbf{P}_n * \lambda_n} - s^{\tilde{\theta}_n}\mathbf{P}_{n,s\mathbf{P}_n * \lambda_n}](t)$ $t \in \mathbf{R}$ where $\lambda_n$ in uniform over $[-a_n, a_n]$ $a_n \to 0$.

The fact that (2.15) is an alternative follows in the same way that theorem 2.3 plus

$\sup_t |s\mathbf{P}_n * \lambda_n - s\mathbf{P})(-\infty, t]| \to 0$ a.s.

Remember that for probabilities measures $\mathbf{P}$ and $\mathbf{Q}$ $\mathbf{P} * \mathbf{Q}$ is defined by

$\mathbf{P} * \mathbf{Q}(A) = \int\int I_{\{x+y\in A\}}\ d\mathbf{P}(x)\ \ d\mathbf{Q}(y) = \int \mathbf{P}(A-y)\ d\mathbf{Q}(y)$.

Hence $|(s\mathbf{P}_n * \lambda_n - s\mathbf{P})(-\infty,t]| = |(1/2a_n)\int_{-a_n}^{a_n} s\mathbf{P}_n(-\infty, t-y] - s\mathbf{P}(-\infty,t]dy| \leq$

$|(1/2a_n)\int_{-a_n}^{a_n} s\mathbf{P}_n(-\infty,t-y] - s\mathbf{P}(-\infty,t-y]dy| + |(1/2a_n)\int_{-a_n}^{a_n} s\mathbf{P}(-\infty,t-y] - s\mathbf{P}(-\infty,t]dy| \leq$

$||s\mathbf{P}_n - s\mathbf{P}||_\infty + 2\sup_{|s-t|\leq a_n} |\mathbf{F}(t) - \mathbf{F}(s)|$

Thus $||s\mathbf{P}_n * \lambda_n - s\mathbf{P}||_\infty \leq ||s\mathbf{P}_n - s\mathbf{P}||_\infty + 2\sup_{|s-t|\leq a_n} |\mathbf{F}(t) - \mathbf{F}(s)| \to 0$ a.s.

Now the analogous of the part (iii) of the theorem 3.2 is that if (2.18) holds jointly with

(2.19) $\sqrt{n}(\theta(\mathbf{P}_{n,s\mathbf{P}_n * \lambda_n}) - \theta(\mathbf{P}_n * \lambda_n)) \to_w \tilde{\theta}(\mathbf{P})$ a.s. then

$\sqrt{n}(\mathbf{P}_{n,s\mathbf{P}_n*\lambda_n} - s\mathbf{P}_{n,s\mathbf{P}_n*\lambda_n}) \rightarrow_w 2^{-1}\mathbf{U}(s\mathbf{F}(\cdot)) + 2^{-1}\mathbf{U}(1 - s\mathbf{F}(\cdot)) + \tilde{\theta}(s\mathbf{P})(sf)(\cdot)$ a.s.

**Corollary 2.5.** Let $\pi \in \Pi$, let $\theta$ be a weak bootstrap consistent parameter of location of symmetry for $\Pi$, and suppose the c.d.f of $\|Z\|_\infty := \|2^{-1}U(s\mathbf{F}(t)) + 2^{-1}U(1 - s\mathbf{F}(t)) + (s^\theta f)(t)\tilde{\theta}(s\mathbf{P})\|_\infty$ is continuous. Let $\Pi_S$ be the set of symmetric probability measures in $\Pi$. Let $t_{n,\alpha}(\omega)$ :=inf$\{t : \hat{\mathbf{Pr}}\{\|n^{1/2}(\mathbf{F}_{n,s\mathbf{P}_n(\omega)} - s\mathbf{F}_{n,s\mathbf{P}_n(\omega)}\|_\infty \geq t\} \geq \alpha\}$.

Consider the test

(2.20) $H_0 : P \in \Pi_S$ versus $H_1 : P \in \Pi - \Pi_S$ with rejection region $\|n^{1/2}(F_n - sF_n)\|_\infty \geq t_{n,\alpha}$

then,

(2.21) $Pr\{\|n^{1/2}(F_n - sF_n)\|_\infty \geq t_{n,\alpha}|H_0\} \rightarrow \alpha$ whereas

(2.22) $Pr\{\|n^{1/2}(F_n - sF_n)\|_\infty \geq t_{n,\alpha}|H_1\} \rightarrow 1$.

Moreover if $\theta$ is strong bootstrap consistent, $t_\alpha$ is defined as $\mathbf{P}\{\|Z\|_\infty > t_\alpha\} = \alpha$ and the c.d.f. of $\|Z\|_\infty$ is strictly increasing at $t_\alpha$ , then $t_{n,\alpha}(\omega) \rightarrow t_\alpha$ a.s.

**Proof.** The first limit (2.21) is a direct consequence of Theorem 2.3. As far as (2.22) just note that if $\mathbf{P}$ is not symmetric, there is $t_0$ such that $\mathbf{F}(t_0) \neq 1 - \mathbf{F}(2\theta(\mathbf{P}) - t_0)$, so that $\|\mathbf{F} - s\mathbf{F}\|_\infty > 0$ Now, $\|\mathbf{F} - s\mathbf{F}\|_\infty \leq \|\mathbf{F} - \mathbf{F}_n\|_\infty + \|\mathbf{F}_n - s\mathbf{F}_n\|_\infty + \|s\mathbf{F}_n - s\mathbf{F}\|_\infty$ and $\|\mathbf{F} - \mathbf{F}_n\|_\infty \rightarrow 0$ a.s., $\|s\mathbf{F} - s\mathbf{F}_n\|_\infty \rightarrow 0$ a.s.,(by Glivenko-Cantelli and a.s. convergence of $\theta_n$ to $\theta$). Therefore, lim $\inf_{n\rightarrow\infty}\|\mathbf{F}_n - s\mathbf{F}_n\|_\infty \geq \|\mathbf{F} - s\mathbf{F}\|_\infty > 0$ a.s., and (2.22) now follows from theorem 2.3. (iii). □

Consider now the local alternatives $\mathbf{Q}_n = \mathbf{P} + n^{-\lambda}\Delta$ $0 < \lambda \leq 1/2$ where $\mathbf{Q}_n$ , $\mathbf{P} \in \Pi$ $\mathbf{P}$ is symmetric but $\mathbf{Q}_n$ is not. Assume $\mathbf{P}$ and $\mathbf{Q}_n$ have bounded and uniformly continuous densities. Suppose that we have

(2.23) $n^{\lambda}(\theta(\mathbf{Q}_n) - \theta(\mathbf{P})) \to \eta(\mathbf{P}, \Delta, \lambda) \quad 0 < \lambda \leq 1/2$

(2.24) $\theta(\mathbf{P}_{n,\mathbf{Q}_n}) \to \theta(\mathbf{P})$ a.s.

For some $0 < \alpha(\lambda) \leq 1/2$ with $\alpha(1/2) = 1/2$ the following limits holds simultaneously

(2.25) $n^{\alpha(\lambda)}(\theta(\mathbf{P}_{n,\mathbf{Q}_n}) - \theta(\mathbf{Q}_n)) \to_w \tilde{\theta}(\mathbf{P}, \Delta, \lambda)$ and

(2.25)' $n^{1/2}(\mathbf{P}_{n,\mathbf{Q}_n} - \mathbf{Q}_n) \to_w U(\mathbf{F}(t))$

Also the following limits holds simultaneously a.s.

(2.26) $n^{\lambda}(\theta(\mathbf{P}_{n,s\mathbf{P}_{n,\mathbf{Q}_n}}) - \theta(\mathbf{P}_{n,\mathbf{Q}_n})) \to_w \tilde{\theta}(\mathbf{P})$.

(2.26)' $n^{1/2}(\mathbf{P}_{n,s\mathbf{P}_{n,\mathbf{Q}_n}} - s\mathbf{P}_{n,\mathbf{Q}_n}) \to_w U(\mathbf{F}(t))$.

The notation is as follows given $(\xi_n)$ i.i.d. v.r. with uniform distribution on [0,1] the sample of size n that we take is $\mathbf{F}_{Q_n}^{-1}(\xi_1), \ldots, \mathbf{F}_{Q_n}^{-1}(\xi_n)$.

**Proposition 2.6.** The following holds under the asumptions of the previous paragraph:

(i) for $\lambda = 1/2$ $n^{1/2}(\mathbf{P}_{n,\mathbf{Q}_n} - s\mathbf{P}_{n,\mathbf{Q}_n})(t) \to 2^{-1}U(s\mathbf{F}(t)) + 2^{-1}U(1 - s\mathbf{F}(t)) + 2^{-1}d(\Delta, t) + [\eta(\mathbf{P}, \Delta, 1/2) + \tilde{\theta}(\mathbf{P}, \Delta, 1/2)]f(t)$ where $d(\Delta, t) = \Delta(-\infty, t] - \Delta[2\theta - t, \infty)$

(ii) for $0 < \lambda < 1/2$ assuming

If $\alpha(\lambda) < \lambda$ and $\tilde{\theta}(\mathbf{P}, \Delta, \lambda) \neq 0$ a.s.then $\|n^{1/2}(\mathbf{P}_{n,\mathbf{Q}_n} - s\mathbf{P}_{n,\mathbf{Q}_n})\|_\infty \to \infty$ a.s.

If $\alpha(\lambda) = \lambda$ and $\|(\eta(\mathbf{P}, \Delta, \lambda) + \tilde{\theta}(\mathbf{P}, \Delta, \lambda))f(t) + d(\Delta, t)\|_\infty \neq 0$ then $\|n^{1/2}(\mathbf{P}_{n,\mathbf{Q}_n} - s\mathbf{P}_{n,\mathbf{Q}_n})\|_\infty \to \infty$ in pr.

If $\alpha(\lambda) > \lambda$ and $\|2^{-1}d(\Delta, t) + \eta(\mathbf{P}, \Delta, \lambda)f(t)\|_\infty \neq 0$ then

$\|n^{1/2}(\mathbf{P}_{n,\mathbf{Q}_n} - s\mathbf{P}_{n,\mathbf{Q}_n})\|_\infty \to \infty$ a.s.

(iii) the process $n^{1/2}(\mathbf{P}_{n,s\mathbf{P}_{n,\mathbf{Q}_n}} - s\mathbf{P}_{n,s\mathbf{P}_{n,\mathbf{Q}_n}}) \to 2^{-1}U(s\mathbf{F}(t)) + 2^{-1}U(1 - s\mathbf{F}(t)) + (s^{\theta}f)(t)\tilde{\theta}(s\mathbf{P})$ converges weakly in $l_\infty(\mathbf{R})$ a.s to the limit procesess of theorem 2.3.

As a consequence, the tests described in corollary 2.5 have asymptotic power 1 against the local alternatives $\mathbf{P} + n^{-\lambda}\Delta$ if $\lambda < 1/2$ and asymptotic power $\mathbf{P}\{\|Z_\Delta\|_\infty > t_\alpha\}$ if $\lambda = 1/2$ where $\mathbf{P}\{\|Z\|_\infty > t_\alpha\} = \alpha$, $Z_\Delta$ is the limit in prosition 2.6 and Z is the limit in theorem 2.3.

We omit the proof of proposition 2.6. It suffices to say that it follows very closely that of theorem 2.3.

### 3. Test based on the Shuster-Narvarte parameter of location of symmetry.

Remember that in section 1 we say that a good measure of how symmetric is a c.d.f. $\mathbf{F}$ is $\inf\{\|\mathbf{F} - \mathbf{G}\|_\infty : \mathbf{G}$ is a symmetric c.d.f.$\}$. First note the following

**Lemma 3.1.** (Schuster (1987)) Given a probability measure $\mathbf{P}$

(3.1) $\inf\{\|\mathbf{P} - \mathbf{Q}\|_\infty : \mathbf{Q}$ is a probability measure symmetric respect to $\theta\} = \|\mathbf{P} - s^\theta\mathbf{P}\|_\infty$

where $s^\theta\mathbf{P}(A) = \frac{1}{2}(\, \mathbf{P}(A) + \mathbf{P}(2\theta - A)\,)$ and $A \in \mathcal{B}$.

**Proof.** Note $\|\mathbf{P} - \mathbf{Q}\|_\infty = \sup_{x<\theta}|(\mathbf{P} - \mathbf{Q})(-\infty, x]| \vee \sup_{\theta\leq x}|(\mathbf{P} - \mathbf{Q})(-\infty, x]|$ and that $\sup_{\theta\leq x}|(\mathbf{P} - \mathbf{Q})(-\infty, x]| = \sup_{\theta<x}|(\mathbf{P} - \mathbf{Q})(-\infty, x]| =$ $\sup_{\theta<x}|(\mathbf{P} - \mathbf{Q})[x, \infty))| = \sup_{\theta>x}|(\mathbf{P} - \mathbf{Q})[2\theta - x, \infty))|$.

Hence $\|\mathbf{P} - \mathbf{Q}\|_\infty = \sup_{x<\theta}\ |(\mathbf{P} - \mathbf{Q})(-\infty, x]| \vee |(\mathbf{P}_n - \mathbf{Q})[2\theta - x, \infty)|$.

Analogously $\|\mathbf{P} - s^\theta\mathbf{P}\|_\infty = \sup_{x<\theta}|(\mathbf{P}(2^{-1}I_{(-\infty,x]} - 2^{-1}I_{[2\theta-x,\infty)}|$.

Note that for $a, b, x \in \mathbf{R}$ $|a - x| \vee |b - x| \leq |(a - b)/2|$. So $|(\mathbf{P} - \mathbf{Q})(-\infty, x]| \vee |(\mathbf{P} - \mathbf{Q})[2\theta - x, \infty))| \leq |\mathbf{P}(2^{-1}I_{(-\infty,x]} - 2^{-1}I_{[2\theta-x,\infty)}|$. Therefore $\|\mathbf{P} - \mathbf{Q}\| \leq \|\mathbf{P} - s^\theta\mathbf{P}\|$. □

Now the estimator of the center of symmetry that we take is the $\theta$ that minimizes $\|F - s^\theta F\|_\infty$

Given $\mathbf{P} \in \mathcal{P}(\mathbf{R})$ and $a \in \mathbf{R}$, let

(3.2) $D^+(a) = D^+(a, \mathbf{P}) = \sup_t \mathbf{P}(I_{(\infty,t]} - I_{[2a-t,\infty)})$

$$D^-(a) = D^-(a,\mathbf{P}) = \sup_t \mathbf{P}(I_{[2a-t,\infty)} - I_{(\infty,t]}).$$

Note that

$$(3.3) \quad \mathbf{P}(I_{(-\infty,a]} - I_{[a,\infty)}) \leq D^+(a) \leq \mathbf{P}(I_{(-\infty,a]}) \quad \mathbf{P}(I_{[a,\infty)} - I_{(-\infty,a]}) \leq D^-(a) \leq \mathbf{P}(I_{[a,\infty)}).$$

So $D^+(a)$ is left-continuous, nondecreasing $D^+(-\infty) = 0 \quad D^+(\infty) = 1$

$D^-(a)$ is right-continuous, nonincreasing $D^-(-\infty) = 1 \quad D^+(\infty) = 0$.

We then let

$$(3.4) \quad \theta^* = \theta^*(\mathbf{P}) = \sup\{a : D^+(a) < D^-(a)\}$$

$$\theta^{**} = \theta^{**}(\mathbf{P}) = \sup\{a : D^+(a) > D^-(a)\}$$

$$\theta = \theta(\mathbf{P}) = \tfrac{1}{2}(\theta^* + \theta^{**})$$

From all the facts juts stated it follows that $\theta(\mathbf{P})$ proved that minimizes the function

$$(3.5) \quad D(a,\mathbf{P}) = D^+(a,\mathbf{P}) \vee D^-(a,\mathbf{P}) = \|F - s^a F\|_\infty.$$

We call $\theta(\mathbf{P})$ the Schuster Narvarte parameter.They found $\theta(\mathbf{P}_n)$ for the empirical distribution. Let $y_i$ be the nondecreasing rearrangement fo $x_1, \ldots, x_n$ including repitions if any, let $\theta_r^+ = 2^{-1} min_{r \leq k \leq n}(y_r + y_{n-k+r})$ and $\theta_r^- = 2^{-1} min_{r \leq k \leq n}(y_{k-r+1} + y_{n-k+1})$. If $r_0 = min\{r : \theta_r^- \leq \theta_r^+\}$ then $\theta^*(\mathbf{P}_n) = max(\theta_{r_0}^-, \theta_{r_0-1}^+)$ , $\theta^{**}(\mathbf{P}_n) = max(\theta_{r_0-1}^-, \theta_{r_0}^+)$ and $\|\mathbf{P}_n - s^{\theta_n}\mathbf{P}_n\| = r_0/n$.

For the SN parameter we let $\Pi_{SN}$ be the set of probability measures $\mathbf{P}$ on $\mathbf{R}$ which are absolutely continuous, with uniformly continuous densities on $D_\mathbf{P}$ and such that $D(a,\mathbf{P})$ attains its minimum at a single point (which is therefore $\theta(\mathbf{P})$). This last condition holds even when the median is not unique, but it does not hold in general. For example, it can be seen that if an absolutely continuous c.d.f. has only one flat, or is symmetric with respect to some center then D(a) attains its minimum at a single point. We check this property for symmetric distributions. If $\mathbf{P}$ is symmetric and $\theta$ is its center of symmetry then $D^+(a) > 0$ for $a > \theta$ and $D^+(a) = 0$ for $a \leq \theta$, similarly $D^-(a) > 0$ for $a < \theta$, $D^-(a) = 0$ for $a \geq \theta$. Hence D attains its minimum only at $\theta$.

**Lemma 3.2.** Let now $\mathbf{P} \in \Pi, \theta = \theta(\mathbf{P})$ and $\theta_n = \theta(\mathbf{P}_n)$. We show that

(3.6) $\theta_n \to \theta$ a.s.

**Proof.** By Glivenko-Cantelly, for all $a$. $D^+(a, \mathbf{P}_n) \to D^+(a, \mathbf{P})$ a.s. $D^-(a, \mathbf{P}_n) \to D^-(a, \mathbf{P})$ a.s. For each $\epsilon > 0$, taking $a = \theta + \epsilon$ ( note $\theta = \theta^* = \theta^{**}$) it follows that
$\lim D^+(\theta + \epsilon, \mathbf{P}_n) = D^+(\theta + \epsilon, \mathbf{P}) > D^-(\theta + \epsilon, \mathbf{P}) = \lim D^-(\theta + \epsilon, \mathbf{P}_n)$ so that from some $n(\omega)$ on $D^+(\theta + \epsilon, \mathbf{P}_n) > D^-(\theta + \epsilon, \mathbf{P}_n)$, hence, $\theta + \epsilon \geq \theta^{**}(\mathbf{P}_n) \geq \theta(\mathbf{P}_n)$. Similarly, taking $a = \theta - \epsilon$ gives $\theta - \epsilon \leq \theta^*(\mathbf{P}_n) \leq \theta(\mathbf{P}_n)$ from some n on and (3.6) follows. □

We will use an 'special construction ', it is well known Shorack (1972) and Shorack-Wellner (1986) (p.93) that as a consequence of a theorem of Skorohod, there exist a Brownian bridge U on [0,1] and, for each n i.i.d. uniform on [0,1] random variables $\xi_{n,1}, \ldots, \xi_{n,n}$ such that if
$U_n(t) = n^{-1/2} \sum_{i=1}^{n} (I_{\{\xi_{n,i} \leq t\}} - t)$ $t \in [0,1]$ then

(3.7) $\|U_n - U\|_\infty \to 0$ a.s.

Define $\bar{\mathbf{P}}_n = n^{-1} \sum_{i=1}^{n} \delta_{F^{-1}(\xi_{n,i})}$ then $\mathcal{L}\,(\bar{\mathbf{P}}_n) = \mathcal{L}\,(\mathbf{P}_n)$.

**Lemma 3.3.** For all $\mathbf{P} \in \Pi_{SN}$ symmetric there is a random variable $\tilde{\theta}(\mathbf{P})$ such that the followings limits holds jointly weakly

(3.8) $\sqrt{n}(\mathbf{P}_n - \mathbf{P})(t) \to U(\mathbf{F}(t))$ $\quad \sqrt{n}(\theta_n - \theta) \to \tilde{\theta}(\mathbf{P})$.

**Proof.** Since $\mathcal{L}\,(\mathbf{P}_n) = \mathcal{L}\,(U_n(\mathbf{F}(t))$, it is enough to prove that $\sqrt{n}(\theta(\mathbf{P}_n) - \theta(\mathbf{P})\ ) \to \tilde{\theta}(\mathbf{P})$ a.s. where $\tilde{\theta}(\mathbf{P})$ will be defined later.

We have $D^+(a, \bar{\mathbf{P}}_n) = \sup_t \bar{\mathbf{P}}_n(I_{(\infty,t]} - I_{[2a-t,\infty)})$ and $D^-(a, \bar{\mathbf{P}}_n) = \sup_t \bar{\mathbf{P}}_n(I_{[2a-t,\infty)} - I_{(\infty,t]})$

(3.9) $E_n^+(a) = \sup_t(\, U_n(\mathbf{F}(t)) + U_n(1-\mathbf{F}(t)) + 2af(t)\,)$

(3.9)' $E_n^-(a) = \sup_t(\, -U_n(\mathbf{F}(t)) - U_n(1-\mathbf{F}(t)) - 2af(t)\,)$

and $E^+(a)$ and $E^-(a)$ to defined by the same expressions with $U_n$ replaced by U. We then have $\sqrt{n}\bar{\mathbf{P}}_n(I_{(\infty,t]} - I_{[2\theta+2an^{-1/2}-t,\infty)}) =$ $\sqrt{n}(\bar{\mathbf{P}}_n - \mathbf{P})(I_{(\infty,t]} - I_{[2\theta-t,\infty)}) + \sqrt{n}(\bar{\mathbf{P}}_n - \mathbf{P})(I_{[2\theta-t,\infty)} - I_{[2\theta+2an^{-1/2}-t,\infty)}) + \sqrt{n}(\mathbf{F}(t) - \mathbf{F}(t - 2an^{-1/2}))$

Therefore,

(3.10) $|\sqrt{n}D^+(\theta + an^{-1/2}, \bar{\mathbf{P}}_n) - E_n^+(a)| \leq$ $\sqrt{n}\|(\mathbf{P}_n - \mathbf{P})I_{\{t\}}\|_\infty + \|U_n(\mathbf{F}(2\theta-t)) - U_n(\mathbf{F}(2\theta+2an^{-1/2}-t))\|_\infty + \|\sqrt{n}(\mathbf{F}(t) - \mathbf{F}(t-2an^{-1/2})) - 2af(t)\|_\infty$.

Now by differentiability properties of **F** and (3.10) together with the uniform continuity of the sample paths of U, $|\sqrt{n}D^+(\theta + an^{-1/2}, \bar{\mathbf{P}}_n) - E_n^+(a)| \to 0$ a.s. for all $a \in \mathbf{R}$

Since $|E_n^+(a) - E^+(a)| \leq \|(U_n - U)(\mathbf{F}(t)) - (U_n - U)(\mathbf{F}(2\theta - t))\|_\infty \to 0$ a.s. We conclude, using monotonicity and continuity of $E^+(a)$ that

(3.11) $\sqrt{n}D^+(\theta + an^{-1/2}, \bar{\mathbf{P}}_n) \to E^+(a)$ for all $a \in \mathbf{R}$ a.s. similarly

$\sqrt{n}D^-(\theta + an^{-1/2}, \bar{\mathbf{P}}_n) \to E^-(a)$ for all $a \in \mathbf{R}$ a.s.

It is shown in Rao-Schuster-Litteli (1975) ( proof of lemma 1) that for each a $\mathbf{P}\{E^+(a) = E^-(a)\} = 0$. Since $E^+$ and $E^-$ are continuous,$E^+ \nearrow$ $E^- \searrow$ $E^+(-\infty) = E^-(\infty) = \infty$ $E^+(\infty) = E^-(-\infty) = -\infty$, it follows that $E^+(r) \neq \mathbf{E}^-(\mathbf{r})$ for all r rational, $\omega - a.s.$ Therefore $\mathbf{E}^+(a) = \mathbf{E}^-(a)$ has a.s. a unique solution $\tilde{\theta}(\mathbf{P})$ which is clearly a random variable.

Given $\epsilon > 0$ $E^-(\tilde{\theta} + \epsilon) < E^+(\theta + \epsilon)$ a.s. so that by (3.10) eventually a.s. $D_-(\theta + (\tilde{\theta} + \epsilon)n^{-1/2}, \bar{\mathbf{P}}_n) < D_+(\theta + (\tilde{\theta} + \epsilon)n^{-1/2}, \bar{\mathbf{P}}_n)$ implying $\limsup_{n\to\infty}\sqrt{n}(\theta^{**}(\bar{\mathbf{P}}_n) - \theta) \leq \tilde{\theta}$ a.s. Similarly using $a = \tilde{\theta} - \epsilon$, we obtain $\liminf_{n\to\infty}\sqrt{n}(\theta^{**}(\bar{\mathbf{P}}_n) - \theta) \geq \tilde{\theta}$ a.s. Therefore $\lim_{n\to\infty}\sqrt{n}(\theta^{**}(\bar{\mathbf{P}}_n) - \theta) = \tilde{\theta}$ a.s. □

Rao-Schuster-Littell (1975, theorem 4) proved weak convergence of $n^{1/2}(\theta(\mathbf{P}_n)-\theta)$ and obtain that law of $\tilde{\theta}$ which is not normal

(3.12) $\mathbf{P}\{\tilde{\theta} \le t\} = \mathbf{P}\{sup_{0<u\le 1}(W(u) + 2tf(F^{-1}(u))+ \inf_{0<u\le 1/2}(W(u) + 2tf(F^{-1}(u)) \ge 0\}$

where W(u) is a brownian motion.

The former proof is somehow similar to theirs. The nice thing it is that it allows bootstrapping.

**Lemma 3.4.** For all $\mathbf{P} \in \Pi$ such that the Hölder condition (2.5) holds there is a random variable $\tilde{\theta}(s\mathbf{P})$ such that $n^{1/2}(\mathbf{P}_{n,s\mathbf{P}_n} - s\mathbf{P}_n) \xrightarrow{\mathcal{L}} U(s\mathbf{F}(t))$ and

$n^{1/2}(\theta(\mathbf{P}_{n,s\mathbf{P}_n}) - \theta(\mathbf{P}_n)) \xrightarrow{\mathcal{L}} \tilde{\theta}(s\mathbf{P})$ hold simultaneously a.s.

**Proof.** We now use the special construcion on $\omega$. We use $\bar{\mathbf{P}}_{n,s\mathbf{P}_n} = n^{-1}\sum_{i=1}^{n} \delta_{(s\mathbf{F}_n)^{-1}(\xi_{n,i})}$

$\bar{E}_n^+(a)$ $=\sup_t(U_n(s\mathbf{F}(t) + U_n(1 - s\mathbf{F}(t) + 2a\,sf(t))$  $\bar{E}_n^-(a)$ $=\sup_t(-U_n(s\mathbf{F}(t) - U_n(1 - s\mathbf{F}(t) - 2a\,sf(t))$

As before $D^+(\theta_n + an^{-1/2}, s\mathbf{P}_n) = sup_t(\sqrt{n}\bar{\mathbf{P}}_{n,s\mathbf{P}_n}(I_{(-\infty,t]} - I_{[2\theta_n+2an^{-1/2}-t,\infty)})) =$ $\sup_t(\sqrt{n}(\bar{\mathbf{P}}_{n,s\mathbf{P}_n} - s\mathbf{P}_n)(I_{(-\infty,t]} - I_{[2\theta_n+2an^{-1/2}-t,\infty)})) + \sqrt{n}s\mathbf{P}_n(I_{(-\infty,t]} - I_{[2\theta_n+2an^{-1/2}-t,\infty)}))$

So $|\sqrt{n}\bar{D}^+(\theta + an^{-1/2}, \bar{\mathbf{P}}_n) - E_n^+(a, s\mathbf{P}_n)| \le$ $\sup_t|(\sqrt{n}\bar{\mathbf{P}}_{n,s\mathbf{P}_n} - s\mathbf{P}_n(I_{(-\infty,t]} - I_{[2\theta_n+2an^{-1/2}-t,\infty)}))| + sup_t|\sqrt{n}s\mathbf{P}_n(I_{(-\infty,t]} - I_{[2\theta_n+2an^{-1/2}-t,\infty)}))|$ =I+II.

I$=\sup_t|U_n(\mathbf{F}_n(2\theta_n - t)) - U_n(\mathbf{F}_n(2\theta_n - t + 2an^{-1/2})|$ because $U_n \to U$ and $\mathbf{F}_n \to \mathbf{F}$ uniformly in t.

Also $\sqrt{n}s\mathbf{P}_n(I_{(-\infty,t]} - I_{2\theta_n+2an^{-1/2}-t,\infty)}) - 2asf(t) =$ $\sqrt{n}(\mathbf{P}_n(2^{-1}I_{(-\infty,t]} + 2^{-1}I_{[2\theta_n-t,\infty)} - 2^{-1}I_{[2\theta_n+2an^{-1/2}-t,\infty)} - 2^{-1}I_{(-\infty,t-2an^{-1/2}]}) - 2asf(t) =$ $\sqrt{n}(\mathbf{P}_n - \mathbf{P})(2^{-1}I_{(-\infty,t]} + 2^{-1}I_{[2\theta_n-t,\infty)} - 2^{-1}I_{[2\theta_n+2an^{-1/2}-t,\infty)} - 2^{-1}I_{(-\infty,t-2an^{-1/2}]})+$

$\sqrt{n}\mathbf{P}(2^{-1}I_{(-\infty,t]}+2^{-1}I_{[2\theta_n-t,\infty)}-2^{-1}I_{[2\theta_n+2an^{-1/2}-t,\infty)}-2^{-1}I_{(-\infty,t-2an^{-1/2}]})-2asf(t)=II_a+II_b$.

We obtain that $\omega - a.s.$ for all $a \in \mathbf{R}$ $D^+(\theta_n + an^{-1/2}, s\mathbf{P}_n) \to E^+(a)$ a.s.

The same holds for $D^-$ and therefore proceding as in the proof of (i) it follows that $\omega$ a.s. $n^{1/2}(\theta(\mathbf{P}_{n,s\mathbf{P}_n}) - \theta(\mathbf{P}_n)) \to \tilde{\theta}(s\mathbf{P})$. □

In order to apply Corollary 2.5 we need that if $\mathbf{F}$ is symmetric the distribution of $\|Z\|_\infty$ is continuous and strictly increasing at every $t_\alpha$. Next will show that this happen for every $\mathbf{P}$. First note that $\|Z\|_\infty = \inf_a \sup_t |2^{-1}U(t) + 2^{-1}U(1-t) + af(F^{-1}(t))|$.

Since $W(t) = (U(t) + U(1-t))/\sqrt{2}$ is a Brownian motio $\sqrt{2}\|Z\|_\infty = \inf_a \sup_t |W(t) + af(F^{-1}(t))|$.

Note that $q : C[0,1] \to \mathbf{R}$ defined by $q(x) = \inf_a \sup_t |x(t) + af(F^{-1}(t))|$ is seminorm. By Borell (1974) see (1.13) in Hoffman-Jorgensen-Shepp-Dudley (1979) $\log \mathbf{P}\{q(W) \le t\}$ is concave. So it has a density and it is strictly increasing in $\{t : 0 < \mathbf{P}\{q(W) \le t\} < 1\}$.

Since C[0,1] is separable by the argument in corollary 2.2 in Hoffman-Jorgensen-Shepp-Dudley (1979) $\mathbf{P}\{q(W) \le t\} > 0$ for each $t > 0$. Also by theorem 1.1 in the same paper $\lim_{t\to 0} t^{-2} \log \mathbf{P}\{q(W) > t\} < 0$. So $\mathbf{P}\{q(W) > t\} < 1$ for each t.

Regarding (2.23)-(2.26), we have that (2.23) holds and moreover:

(3.13) $\eta(\mathbf{P}, \Delta, \lambda) = \eta(\mathbf{P}, \Delta)$ $0 < \lambda \le 1/2$ is the unique solution if exists of the equation

(3.14) $\sup_t(2af(t) + d(\Delta, t)) = \sup_t(-2af(t) - d(\Delta, t))$

(3.15) $\alpha(\lambda) = \lambda$ $\tilde{\theta}(\mathbf{P}, \Delta, 1/2)$ is the unique solution in a of

(3.16) $\sup_t(\mathbf{U}(\mathbf{F}(t)) + \mathbf{U}(1-\mathbf{F}(t)) + 2af(t) + 2\eta(\mathbf{P}, \Delta, \lambda)\ f(t) + d(\Delta, t)) =$

$\sup_t(-\mathbf{U}(\mathbf{F}(t)) - \mathbf{U}(1-\mathbf{F}(t)) - 2af(t) - 2\eta(\mathbf{P}, \Delta, \lambda)\ f(t) - d(\Delta, t))$

and for $0 < \lambda < 1/2$

(3.17) $\tilde{\theta}(\mathbf{P}, \Delta, \lambda)$ is the unique solution if exists of the equation

$\sup_t(2af(t) + 2\eta(\mathbf{P}, \Delta, \lambda)\, f(t) + d(\Delta, t)) = \sup_t(-2af(t) - 2\eta(\mathbf{P}, \Delta, \lambda)\, f(t) - d(\Delta, t))$.

In conclusion.

**Theorem 3.4.** The SN parameter of location of symmetry is strong bootstrap consistent with $\tilde{\theta}(\mathbf{P})$ given as the a.s. unique solution of $E^+(a) = E^-(a)$. And it also satisties (2.23)-(2.26) with variables as defined in (3.13)-(3.17). Therefore theorem 2.3 and corollary 2.5 holds for $\sqrt{n}(\mathbf{P}_n - s\mathbf{P}_n)$.

It can be checked that the map $\mathbf{P} \longrightarrow s\mathbf{P}$ is not smoot for the SN parameter, but this follows also from the fact that Z(t) is not gaussian in general. In the case that **F** has a uniform density $\mathbf{P}\{\tilde{\theta} \leq t\} = \mathbf{P}\{\sup_{0 \leq u \leq 1} W(u) + \inf_{0 \leq u \leq 1} W(u) \leq 4t\}$.

Now the joint distribution of $m = \inf_{0 \leq t \leq 1} W(t)$ $M = \sup_{0 \leq t \leq 1} W(t)$ is well kown. See e.g. Billinsgley (1969) page 79. After a few elemental computations we get that

(3.18) $\mathbf{P}\{\tilde{\theta} \leq t\} = \sum_{k=1}^{\infty}(-1)^{k+1}((2k-1)^{-1} + (2k+1)^{-1})\Phi(4tk)$ where $\Phi$ is the c.d.f. of a normal distribution.

Note M+m has a symmetric distribution. So if $\tilde{\theta}$ were normal variable , it would be a centered normal variable.Its tail is asymptotically like the tail of $\Phi(4t)$. Hence if it were normal, it would have variance 4. But since the series (3.18) is an alternate series

$\mathbf{P}\{\tilde{\theta} \leq t\} \geq (4/3)\Phi(4t) - (8/15)\Phi(8t) > \Phi(4t)$ for t close to $-\infty$. So $\tilde{\theta}$ is not a normal variable.

## 4. Test based on the median.

The median can be studied in an analogous way. The relevant variable for the median m with m =m(P) is D(a,P)=F(a).Given a probability measure **P** when we talk about the median of **P** we refire to

(4.1) $m^{*} = m^{*}(\mathbf{P}) = \sup\{a : \mathbf{F}(a) < 1/2\}$.

$m^{**} = m^{**}(\mathbf{P}) = \inf\{a : \mathbf{F}(a) > 1/2\}$.

$m = m(\mathbf{P}) = \frac{1}{2}(m^{*}(\mathbf{P}) + m^{**}(\mathbf{P}))$

**Lemma 4.1.** If $\mathbf{F}(t) < 1/2$ for $t < m$ and $\mathbf{F}(t) > 1/2$ for $t > m$ then $m(\mathbf{P}_n) \to m(\mathbf{P})$ a.s.

**Proof.** By Glivenko-Cantelli for each a $\mathbf{F}_n(a) \to \mathbf{F}(a)$. For each $\epsilon > 0$ $\mathbf{F}(m+\epsilon) > 1/2 > \mathbf{F}(m-\epsilon)$. Hence for n large $\mathbf{F}_n(m+\epsilon) > 1/2 > \mathbf{F}_n(m-\epsilon)$. Thus $m+\epsilon \geq m(\mathbf{P}_n) \geq m-\epsilon$ . □

**Lemma 4.2.** If $f(m) \neq 0$ then the following limits holds simultaneously $\sqrt{n}(m(\mathbf{P}_n) - m(\mathbf{P})) \xrightarrow{\mathcal{L}} -U(1/2)f(m)^{-1}$ and $\sqrt{n}(\mathbf{P}_n - \mathbf{P}) \xrightarrow{\mathcal{L}} U(\mathbf{F}(t))$

**Proof.** Note that in the construction $\sqrt{n}[\bar{\mathbf{P}}_n(m + an^{-1/2}) - 1/2] \to af(m) + U(1/2)$. □

**Lemma 4.3.** If $f(m) \neq 0$ and $h(\delta) := \sup_{|t-s| \leq \delta} |\mathbf{F}(t) - \mathbf{F}(s)| \ln \ \ln(1/|t-s|) \to 0$ as $\delta \to 0$ holds then the following limits holds simultaneously a.s. $\sqrt{n}(m(\mathbf{P}_{n\,s\mathbf{P}_n}) - m(\mathbf{P}_n)) \xrightarrow{\mathcal{L}} -U(1/2)f(m)^{-1}$ $\sqrt{n}(\mathbf{P}_{n\,s\mathbf{P}_n} - s\mathbf{P}_n) \xrightarrow{\mathcal{L}} U(s\mathbf{F}(t))$.

**Proof.** $\sqrt{n}(\mathbf{P}_{n\,s\mathbf{P}_n}(m(\mathbf{P}_n) + an^{-1/2}) - 1/2) =$

$$\sqrt{n}(\mathbf{P}_{n\,s\mathbf{P}_n} - s\mathbf{P}_n)(m(\mathbf{P}_n) + an^{-1/2}) + \sqrt{n}[s\mathbf{P}_n(m(\mathbf{P}_n) + an^{-1/2}) - 1/2)] = I + II.$$

Now $I \to U(1/2)$.

$II = \sqrt{n}[\mathbf{P}_n(2^{-1}I_{(-\infty,m(\mathbf{P}_n)+an^{-1/2}]} + 2^{-1}I_{[m(\mathbf{P}_n)-an^{-1/2},\infty)}) - 1/2] =$

$\sqrt{n}\mathbf{P}_n(2^{-1}I_{(-\infty,m(\mathbf{P}_n)+an^{-1/2}]} - 2^{-1}I_{(-\infty,m(\mathbf{P}_n)-an^{-1/2})}) =$

$\sqrt{n}(\mathbf{P}_n - \mathbf{P})(2^{-1}I_{(-\infty,m(\mathbf{P}_n)+an^{-1/2}]} - 2^{-1}I_{(-\infty,m(\mathbf{P}_n)-an^{-1/2})}))+$

$\sqrt{n}\mathbf{P}(2^{-1}I_{(-\infty,m(\mathbf{P}_n)+an^{-1/2}]} - 2^{-1}I_{(\infty,m(\mathbf{P}_n)-an^{-1/2})}) = II_a + II_b.$

$\sup_t|II_a| \to 0$ a.s. because (2.17).

$\sup_t|II_b - af(t)| \to 0$ a.s. because f(t) is uniformly differentiable.

Hence in $l_\infty(\mathbf{R})$ $\sqrt{n}[\mathbf{F}_{n,s\mathbf{P}_n}(m(\mathbf{P}_n) + an^{-1/2}) - 1/2] \to U(1/2) + af(t)$. □

The same arguments that for the The Schuster-Narvarte estimator holds here for $||Z||_\infty$. Now $Z(t) = U(\mathbf{F}(t)) + 2^{-1}U(1 - s\mathbf{F}(t)) - f(t)U(1/2)f(m)^{-1}$ is gaussian process and $q : C[0,1] \to \mathbf{R}$ defined by $q(x) = \sup_t|x(t)|$ is a norm.

Call $\Pi_m$ the set of probability measures on $\mathbf{R}$ which are absolutely continuous with bounded uniformly continuous derivatives on $\mathbf{D_P}$ satisfying (2.5) and $f(m(\mathbf{P})) \neq 0$. As far as the local alternatives by computations similar to the former computations $\alpha(\lambda) = \lambda$

$\eta(\mathbf{P}, \Delta, \lambda) = -\mathbf{F}_\Delta(m)(f(m))^{-1}$ for $0 < \lambda \leq 1/2$

$\tilde{\theta}(\mathbf{P}, \Delta, 1/2) = -(U(1/2) + \mathbf{F}_\Delta(m))f(m)^{-1}$

$\tilde{\theta}(\mathbf{P}, \Delta, \lambda) = -\mathbf{F}_\Delta(m)f(m)^{-1}$ for $0 < \lambda < 1/2$.

Therefore $Z(t) = 2^{-1}U(s\mathbf{F}(t)) + 2^{-1}U(1 - s\mathbf{F}(t)) - U(1/2)sf(t)f(m)^{-1}$ is centered gaussian process and $Z_\Delta(t) = Z(t) + 2^{-1}d(\Delta, t) - \mathbf{F}_\Delta(m)f(t)f(m)^{-1}$ is the limit in the proposition 2.6 (iii).

**Theorem 4.1.** The median is a strong bootstrap consistent for $\Pi_m$ and satisties (2.23)-(2.26) with the limiting variables just especified and the limit (ii) in Definition 2.3. holds also if $\mathbf{P}$ is not symmetric.

### 5. Test based on the Hodges-Lehmman parameter.

The analysis for Schuster-Narvarte estimator can be done similarly for other distances. Consider now the distance

(5.1) $d(\mathbf{P},\mathbf{Q}) = \int_{-\infty}^{\infty}(\mathbf{P}(-\infty,x] - \mathbf{Q}(-\infty,x])^2\,dx.$

**Lemma 5.1.** Given a probability measure $\mathbf{P}$

$d(\mathbf{P},s^{\theta}\mathbf{P})$ =inf$\{d(\mathbf{P},\mathbf{Q}) : \mathbf{Q}$ is symmetric respect to $\theta\}$.

Proof. If $\mathbf{Q}$ is symmetric repect to $\theta$ $d(\mathbf{P},\mathbf{Q}) = \int_{-\infty}^{\infty}(\mathbf{P}(-\infty,x+\theta] - \mathbf{Q}(-\infty,x+\theta])^2\,dx =$ $\int_{-\infty}^{0}(\mathbf{P}(-\infty,x+\theta] - \mathbf{Q}(-\infty,x+\theta])^2 + (\mathbf{P}(-\infty,\theta-x] - \mathbf{Q}(-\infty,\theta-x])^2\,dx =$ $\int_{-\infty}^{0}(\mathbf{P}(-\infty,x+\theta] - \mathbf{Q}(-\infty,x+\theta])^2 + (\mathbf{P}(-\infty,\theta-x] - 1 + \mathbf{Q}(-\infty,\theta+x])^2\,dx.$

For each x we have an expresion of the type $(a-M)^2 + (b-M)^2$ that is minimized at $M = (a+b)/2$ i.e. if $\mathbf{Q}(-\infty,\theta+x] = (\mathbf{P}(-\infty,x+\theta] + \mathbf{P}[\theta-x,\infty)/2$. □

In Fine (1966) it is shown that

$d(\mathbf{P},\mathbf{Q}) = \int_{-\infty}^{\infty}(\mathbf{P}(-\infty,x]-\mathbf{Q}(-\infty,x])^2\,dx = \int\int |2\theta-s-t|-|s-t|\,d\mathbf{P}(s)d\mathbf{P}(t)$. So $d(\mathbf{P},s^{\theta}\mathbf{P})$ is minimized in a median of $(\mathbf{X}+\mathbf{X}')/2$ where X' is a copy of X.

Hence a $\theta$ that minimizes $d(\mathbf{P},s^{\theta}\mathbf{P})$ is defined as follows. Define

(5.2) $D(\theta) = \int_{-\infty}^{\infty}\mathbf{F}(2\theta-x)\,d\mathbf{F}(x).$

Define $\theta^{*} = \theta^{*}(\mathbf{P})$ =sup$\{\theta : D(\theta) < 1/2\}$ , $\theta^{**} = \theta^{**}(\mathbf{P})$ =inf$\{\theta : D(\theta) > 1/2\}$ and

(5.3) $\theta = \theta(\mathbf{P}) = \frac{1}{2}(\theta^{*}(\mathbf{P}) + \theta^{**}(\mathbf{P}))$.

In particular, if $\mathbf{X}_i$ are i.i.d. then $\theta(\mathbf{P}_n)$ = center of medians of $\{(\mathbf{X}_i+\mathbf{X}_j)/2\}_{i,j=1}^{n}$.

**Lemma 5.2.** If P is symmetric repect to m then $m = \theta^*(\mathbf{P}) = \theta^{**}(\mathbf{P})$.

**Proof.** Take a point in the support of **F** i.e. $\mathbf{P}\{\mathbf{X} \in (a-\delta, a+\delta)\} > 0$ for each $\delta > 0$. Then $0 < \mathbf{P}\{\mathbf{X} \in (a-\delta, a+\delta) \;\; \mathbf{X}' \in (2m-a-\delta, 2m-a+\delta)\} \leq \mathbf{P}\{\mathbf{X}+\mathbf{X}' \in (2m-2\delta, 2m-2\delta)\}$. So that by symmetry $\mathbf{P}\{(\mathbf{X}+\mathbf{X}')/2 \leq m-\delta\} = \mathbf{P}\{(\mathbf{X}+\mathbf{X}')/2 \geq m-\delta\} < 1/2$, and therefore $\theta^*, \theta^{**} \in [m-\delta, m+\delta]$ for all $\delta$. □

We will show that $\theta$ satisties Defintion 2.2 for the class $\Pi_{HL}$ of absolutely continuous probability measures **P** on $D_{\mathbf{P}}$ such that $\theta^*(\mathbf{P}) = \theta^{**}(\mathbf{P})$.

**Lemma 5.3.** If $\mathbf{P} \in \Pi_{HL}$ then $\theta(\mathbf{P}_n) \to \theta(\mathbf{P})$.

**Proof.** We note

$D(a, \mathbf{P}_n) = \int_{-\infty}^{\infty} \mathbf{F}_n(2a-x) - \mathbf{F}(2a-x)\, d\mathbf{F}_n(x) + \int_{-\infty}^{\infty} \mathbf{F}(2a-x)\, d(\mathbf{F}_n - \mathbf{F})(x) + \int_{-\infty}^{\infty} \mathbf{F}(2a-x)\, d\mathbf{F}(x)$.

The first integral tends to 0 a.s. by Glivenko -Cantelli, the second one also tends to 0 a.s. because **F** is bounded and continuous and $\mathbf{P}_n(\omega) \to \mathbf{P}$ weakly a.s., and the third is $D(a, \mathbf{P})$. Since $\mathbf{P} \in \Pi_{HL}$ for any $\epsilon > 0$, $D(\theta+\epsilon, \mathbf{P}) > 1/2 > D(\theta-\epsilon, \mathbf{P})$. Thus for n large $D(\theta+\epsilon, \mathbf{P}_n) > 1/2 > D(\theta-\epsilon, \mathbf{P}_n)$ eventually a.s. implying $\theta+\epsilon \geq \theta_n^{**} \geq \theta_n^* \geq \theta-\epsilon$. □

**Lemma 5.4.** If $\mathbf{P} \in \Pi_{HL}$ and is symmetric then the following limits holds jointly

$\sqrt{n}(\mathbf{P}_{n\,s\mathbf{P}_n}) - s\mathbf{P}_n) \xrightarrow{\mathcal{L}} U(\mathbf{F}(t))$ and

$\sqrt{n}(\theta(\mathbf{P}_n) - \theta(\mathbf{P}) \xrightarrow{\mathcal{L}} -\int_0^1 U(t)\, d(t) / \int_{-\infty}^{\infty} f^2(x)\, dx$.

**Proof.** Using the special construction we must show that

$\sqrt{n}(\theta(\mathbf{P}_n) - \theta(\mathbf{P})) \to -\int_0^1 U(t)\,dt / \int_{-\infty}^{\infty} f^2(x)\,dx$. Note

$\sqrt{n}[D(\theta + an^{-1/2}, \bar{\mathbf{P}}_n) - 1/2] = \sqrt{n}\int_{-\infty}^{\infty} \bar{\mathbf{F}}_n(2\theta + 2an^{-1/2} - x) - \mathbf{F}(2\theta + 2an^{-1/2} - x)\,d\bar{\mathbf{F}}_n(x) +$

$\sqrt{n}\int_{-\infty}^{\infty} \mathbf{F}(2\theta + 2an^{-1/2} - x) - \mathbf{F}(2\theta - x)\,d(\bar{\mathbf{F}}_n(x) + \sqrt{n}\int_{-\infty}^{\infty} \mathbf{F}(2\theta - x)\,d(\bar{\mathbf{F}}_n - \mathbf{F})(x)$.

By 3.7 and the uniformly continuity of the paths of U, he integrand in the first term converges uniformly in x $\omega$ a.s. to $U(\mathbf{F}(2\theta - x))$ so that this term is a.s. asymptotically equivalent to $\int_{-\infty}^{\infty} U(\mathbf{F}(2\theta - x))d\bar{\mathbf{F}}_n(x)$ which since $\bar{\mathbf{P}}_n \to \mathbf{P}$ weakly a.s., converges a.s. to $\int_{-\infty}^{\infty} U(\mathbf{F}(2\theta - x))d\mathbf{F}(x)$.

By uniformly differentiability of $\mathbf{F}$ the second term tends to $2a\int_{-\infty}^{\infty} f(2\theta - x)\,d\mathbf{F}(x)$ a.s.

Integration by parts together with previous arguments show that the third term converges to $\int_{-\infty}^{\infty} U(\mathbf{F}(2\theta - x)\,d\mathbf{F}(x)$ a.s. Thus we have

$\sqrt{n}[D(\theta + an^{-1/2}, \bar{\mathbf{P}}_n) - 1/2] \to 2a\int_{-\infty}^{\infty} f(2\theta - x)\,d\mathbf{F}(x) + 2\int_{-\infty}^{\infty} U(\mathbf{F}(2\theta - x))\,d\mathbf{F}(x)$.

So $\sqrt{n}(\theta(\bar{\mathbf{P}}_n) - \theta) \to -\int_0^1 U(t)\,d(t) / \int_{-\infty}^{\infty} f^2(x)\,dx$ because if $\mathbf{F}$ is symmetric $\mathbf{F}(2\theta - x) = 1 - \mathbf{F}(x)$ and $f(2\theta - x) = f(x)$. □

**Lemma 5.5.** If $\mathbf{P} \in \Pi_{HL}$ then the following limits holds jointly a.s.

$\sqrt{n}(\mathbf{P}_{n,s\mathbf{P}_n} - s\mathbf{P}_n) \xrightarrow{\mathcal{L}} U(s\mathbf{F}(t))$ and $\sqrt{n}(\theta(\mathbf{P}_{n,s\mathbf{P}_n}) - \theta(\mathbf{P}_n)) \xrightarrow{\mathcal{L}} -\int_0^1 U(t)\,dt / \int_{-\infty}^{\infty} sf^2(x)\,dx$.

**Proof.** Using the special construction we must check

$\sqrt{n}(\theta(\bar{\mathbf{P}}_{n,s\mathbf{P}_n}) - \theta(\mathbf{P}_n)) \to -\int_0^1 U(t)\,dt / \int_{-\infty}^{\infty} sf^2(x)\,dx$.

Since $\|s\mathbf{P}_n\{x\}\|_\infty \leq 1/n$ a.s.

(5.4) $\sqrt{n}[D(\theta_n + an^{-1/2}, \bar{\mathbf{P}}_{n,s\mathbf{P}_n} - 1/2] =$

$\sqrt{n}\int_{-\infty}^{\infty} \bar{\mathbf{F}}_{n,s\mathbf{P}_n}(2\theta_n + 2an^{-1/2} - x) - s\mathbf{F}_n(2\theta_n + 2an^{-1/2} - x)\,d\bar{\mathbf{F}}_{n,s\mathbf{P}_n}(x) +$

$$\sqrt{n}\int_{-\infty}^{\infty} s\mathbf{F}_n(2\theta_n + 2an^{-1/2} - x) - s\mathbf{F}_n(2\theta_n - x)\, d\bar{\mathbf{F}}_{n,s\mathbf{P}_n}(x) +$$

$$\sqrt{n}\int_{-\infty}^{\infty} s\mathbf{F}_n(2\theta_n - x)\, d(\bar{\mathbf{F}}_{ns\mathbf{P}_n} - s\mathbf{F}_n)(x) \; + O(n^{-1/2}) = I + II + III + O(n^{-1/2})$$

Since $||s\mathbf{F}_n - s\mathbf{F}||_\infty \to 0$ a.s. it follows that the integrand in I converges uniformly in x to $U(s\mathbf{F}(2\theta - x)) = U(1 - s\mathbf{F}(x))$. For almost every $\omega$ $||\bar{\mathbf{F}}_{n,s\mathbf{P}_n} - s\mathbf{F}_n||_\infty = n^{-1/2}||U_n(s\mathbf{F}_n)||_\infty \xrightarrow{P} 0$. But the same reasons, therefore $\omega$-a.s. $\mathbf{P}_{n,s\mathbf{P}_n} \xrightarrow{\mathcal{L}} s\mathbf{P}$ and we conclude that

(5.5) $\lim(I) = \int_0^1 U(t)\, dt \quad \hat{\mathbf{P}}$ a.s.

For II we note that since $\mathbf{P}_n$ has a.s jumps of size only 1/n

(5.6) $||s\mathbf{F}_n(2\theta_n + 2an^{-1/2} - x) - s\mathbf{F}_n(2\theta_n - x) - 2^{-1}\mathbf{F}(2\theta_n + 2an^{-1/2} - x) + 2^{-1}\mathbf{F}(2\theta_n - x) + 2^{-1}\mathbf{F}(x) + 2^{-1}\mathbf{F}(x - 2an^{-1/2})||_\infty \leq$

$O(n^{-1/2}) + 2^{-1}||\sqrt{n}(\mathbf{P}_n - \mathbf{P})(I_{(-\infty, 2\theta_n + 2an^{-1/2} - x]} - I_{(-\infty, 2\theta_n - x]} + {(-\infty, x]} - I_{(-\infty, x - 2an^{-1/2}]})||_\infty \to 0$ a.s. by (3.17).

Moreover

$||\mathbf{F}(2\theta_n + 2an^{-1/2} - x) - \mathbf{F}(2\theta_n - x) + \mathbf{F}(x) - \mathbf{F}(x - 2an^{-1/2}) - (f(2\theta - x) + f(x))2an^{-1/2}||_\infty \to 0$

a.s. by uniform differentiability of $\mathbf{F}$. Hence using again that $\bar{\mathbf{F}}_{n,s\mathbf{P}_n} \to s\mathbf{P}$ weakly $\omega$ a.s. we obtain

(5.7) $\lim(\mathrm{III}) = \int_0^1 U(t)\, dt.$

So from (5.4)-(5.7) we obtain that

$\sqrt{n}[D(\theta_n + an^{-1/2}, \bar{\mathbf{P}}_{n,s\mathbf{P}_n} - 1/2] \to 2\int_0^1 U(t)\, d(t) + 2a\int_{-\infty}^{\infty} sf^2(x)\, dx$ a.s. □

The consitency conditions (2.23)-(2.26) can be checked using the same techniques to obtain the following $\eta(\mathbf{P}, \Delta, \lambda) = -\int_{-\infty}^{\infty} \mathbf{F}_\Delta(2\theta - x)\, d\mathbf{F}(x) / \int_{-\infty}^{\infty} f^2(x) dx$ and $\alpha(\lambda) = 1/2\; \tilde{\theta}(\mathbf{P}, \Delta, \lambda) = \tilde{\theta}(\mathbf{P})$.

Note that the limiting process is

$Z(t) = 2^{-1}U(s\mathbf{F}(t)) + 2^{-1}U(1 - s\mathbf{F}(t)) - (\int_0^1 U(t)\; dt / \int_{-\infty}^{\infty} (sf(x))^2(x)\; dx)\; sf(t)$ which is a gaussian process because $\int_0^1 U(t)\, dt$ is the limit of its Riemman sums that are jointly gaussian with the process U(t). So the distribution of $||Z||_\infty$ is continuous and strictly increasing in $(0,\infty)$ and we are in business for Corollary 2.5.

# CHAPTER V

# THE BOOTSTRAP OF M-ESTIMATORS

## 1. M-estimators.

This chapter is a extended exposition of the parts of the article by Arcones-Giné (1990b) about the bootstrap of M-estimators.

A very well kown kind of estimators are the maximum likelihood estimators. They are defined in the following way: suppose that we have a sample from family of probability densities $\{f_\theta(x) : \theta \in \Theta\}$, i.e. there is value $\theta_0$ so that the sample has the distribution of $X_1, \ldots, X_n$ i.i.d. r.v. with density $f_{\theta_0}(\cdot)$ then the problem that we try to solve is to estimate to estimate $\theta_0$. The maximum likelihood method consists in given the sample we consider the likelihood as a function of $\theta$ then we take as an estimator of $\theta_0$ the value that maximizes the likelihood. Since the r.v. are i.i.d. the likelihood of a sample is $\prod_{i=1}^n f_\theta(x_i)$. So the maximum likelihood estimator $\hat{\theta}$ is the value of $\theta$ that satisfies

(1.1) $\prod_{i=1}^n f_{\hat{\theta}}(x_i) = \sup_\theta \prod_{i=1}^n f_\theta(x_i)$.

A reference for the asymptotics of this estimators is Cramer (1943).

The M-estimators, which are a variation of the maximum likelihood estimators, are defined in the following way: suppose we have a function $g(x, \theta)$ and $X_1, \ldots, X_n$ are i.i.d. r.v. and we want to estimate $\theta_0$, the value of $\theta$ that maximizes $Pg(\cdot, \theta)$. The M-estimator $\hat{\theta}(x_1, \ldots, x_n)$ corresponding to the "criterion function" g is any $\hat{\theta}$ such that

(1.2) $\mathbf{P}_n g(\cdot, \hat{\theta}) = \sup_\theta \mathbf{P}_n g(\cdot, \theta)$.

This case includes the former case. If $E_{\theta_0}$ denotes expectation when X has density function $f_{\theta_0}(x)$ then for each $\theta$ $E_{\theta_0} \log f_\theta(\cdot) \leq E_{\theta_0} \log f_{\theta_0}(\cdot)$ and there is inequality if and only if $f_\theta = f_{\theta_0}$ a.s. The maximization of $\mathbf{P}_n \, log \, f_\theta(\cdot)$ is equivalent to the maximization of $\prod_{i=1}^n f_\theta(x_i)$.

Another variation is to consider the value $\theta_0$ of $\theta$ satisfying $Ph(\cdot,\theta_0) = 0$ (assuming uniqueness). Then the estimator that we take is the random value $\hat{\theta}$ solving the equation

(1.3) $\mathbf{P}_n h(\cdot,\hat{\theta}) = 0$

(at least approximately since an eaxct solution may not exist).

Under mild conditions the two estimators coincide where $h = \frac{\delta g}{\delta \theta}$

In this chapter we consider the bootstrap of estimators defined as in (1.2) or in (1.3).

## 2. Background on empirical processes.

Now we will explain some results about the theory of empirical processes that we will need later on.

Let $(S,\mathcal{S},\mathcal{P})$ be a probability space and $\mathcal{F} \subset L_2(S,\mathcal{S},\mathcal{P})$ be a class of measurable function such that $F(s) = \sup\{|f(s)| : f \in \mathcal{F}\} < \infty$ for all $s \in S$ and $\sup\{|\mathbf{P}f|;\ f \in \mathcal{F}\} < \infty$. Then $n^{1/2}(P_n - P) \in \mathrm{l}^\infty(\mathcal{F})$ i.e. it is a $\mathrm{l}^\infty(\mathcal{F})$-valued funtion. $\nu_n^P = n^{1/2}(P_n - P)$ is called the empirical process. Unless $\mathcal{F}$ is finite $\mathrm{l}^\infty(\mathcal{F})$ is not separable and $\nu_n^{\mathbf{P}}$ is not measurable so we cannot use the usual definition of weak convergence. The convergence for elements in $\mathrm{l}^\infty(\mathcal{F})$ that is used (following to Hoffmann-Jorgensen) is as follows: if $\{X_n\}_{n=0}^\infty$ are $\mathrm{l}^\infty(\mathcal{F})$-valued random elements and $X_0$ is measurable and has separable support then $X_n \xrightarrow{\mathcal{L}} X_0$ in $\mathrm{l}^\infty(\mathcal{F})$ iff $\mathbf{E}^* H(X_n) \to \mathbf{E}H(X_0)$ for all $H : \mathrm{l}^\infty(\mathcal{F}) \to \mathbf{R}$ bounded and continuous. $\mathbf{E}^*$ stands for outer integral.

In our case by the finite-dimensional central limit theorem there is a centered Gaussian process $G_P$ indexed by $\mathcal{F}$ with the covariance $\mathbf{E}G_P(f)G_P(g) = \mathbf{E}\,(f - Pf)(g - Pg)$ such that for any finite set $J \subset \mathcal{F}$ $(\nu_n(f) : f \in J) \xrightarrow{\mathcal{L}} (G_P(f); f \in J)$. The kind central limit theorem that we consider is as follows. We say that $\mathcal{F}$ is a P-Donsker class if:

(i) $G_P$ has a version with bounded and $\rho$-uniformly continuous trajectories. Where

$\rho(f,g) = Var_P(f-g)$. Then it is measurable and has separable support.

(ii) $\nu_n^P \xrightarrow{\mathcal{L}} G_P$ in $l^\infty(\mathcal{F})$.

The bootstrap of the empirical processes was studied in Giné-Zinn (1990a):

**Theorem 2.1.** (Gine-Zinn (1990a)). Let $\mathcal{F}$ be a measurable class of functions on S. Then

(i) $\mathcal{F}$ is P-Donsker iff $d^*_{BL}(\nu_n^{\mathbf{P}_n}, G_{\mathbf{P}}) \to 0$ in pr.

(If this happens we say that $\mathcal{F}$ is pr-**P**-Donsker).

(ii) $\mathcal{F}$ is P-Donsker and $\mathbf{E}F^2 < \infty$ iff $d^*_{BL}(\nu_n^{P_n}, G_P) \to 0$ a.s.

(If this happens we say that $\mathcal{F}$ is a.s.-**P**-Donsker).

The second result we will need is tan exponential inequality in Alexander (1984). A class of sets $\mathcal{C}$ is VC ( Vapnik-Cervonenkis ) if there is $n < \infty$ such that $\sup_{x \in S^n} \#\{\ \{x_1, \ldots, x_n\} \cap C; C \in \mathcal{C}\ \} < 2^n$. We refer to Dudley (1984) for the study of the properties of VC classes. A class $\mathcal{F}$ of functions is a VC-graph class if the class of sets $\{(x,t) : 0 \leq t \leq f(x)$ or $f(x) \leq t \leq 0, f \in \mathcal{F}\ \}$ is VC. If $\mathcal{F}$ is VC-graph class then the class $\{f - g : f, g \in \mathcal{F}\}$ is called a VC graph difference class.

**Theorem 2.2** (Alexander (1984)) Let $\mathbf{P} \in \mathcal{P}\ (S)$ and let $\mathcal{F}$ be a measurable uniformly bounded VC-graph difference class of functions on S. Let $\alpha = \sup_{f \in \mathcal{F}} Var_{\mathbf{P}} f$, let $M > 0$. There exists a constant k (depending only on certain characteristics of $\mathcal{F}$) such that if

(2.1) $M^2 \geq k\alpha \log(\alpha^{-1})$ , $M \geq k \log n / n^{1/2}$ and $M \leq 4n^{1/2}\alpha$ then

$\mathbf{P}\{\|\nu_n\|_{\mathcal{F}} > M\} \leq 16exp(-M^2/8\alpha r^2)$ where r=$\sup\{|f(x)| : f \in \mathcal{F}\ , x \in S\}$.

We will use the following corollary of this theorem.

**Corollary 2.3.** Let $\mathbf{P} \in \mathcal{P}$ (S) and let $\mathcal{F}$ be a measurable uniformly bounded VC-graph class of functions. Let $\mathcal{F}'_n = \{f - g : f, g \in \mathcal{F}, \rho^2_{\mathbf{P}}(f,g) \le (\log\log n)^{-(2+\delta)}\}$ for some $c > 0$ and $\delta > 0$. Then $(\log\log n)^{1/2}\|\nu_n\|_{\mathcal{F}'_n} \to 0$ a.s.

**Proof.** By Glivenko-Cantelli we need to prove $\sum_{k=3}^{\infty} \mathbf{P}\{\max_{2^k \le n \le 2^{k+1}} (\log\log n)^{1/2}\|\nu_n\|_{\mathcal{F}'_n} > \epsilon\} < \infty$. But using Ottavianni's inequality (e.g. Araujo and Giné (1980), page 111)
$\mathbf{P}\{\max_{2^k \le n \le 2^{k+1}} (\log\log n)^{1/2}\|\nu_n\|_{\mathcal{F}'_n} > \epsilon\} \le \mathbf{P}\{\|\nu_{2^{k+1}}\|_{\mathcal{F}'_{2^k}} > \epsilon 2^{-3/2}(\log k)^{-1/2}\}$
via the Borel-Cantelli lemma, the result follows. □

## 3. The a.s. bootstrap of M-estimators.

In this section we prove an a.s. bootstrap CLT for M-estimators under conditions close to the non-standard conditions of Huber (1967) and those of Pollard (1985). The proofs are based on methods from these two papers, the bootstrap of empirical measures in Giné and Zinn (1990a), Alexander's (1984) exponential bounds and the "square root trick" of Le Cam in the version of Giné and Zinn (1984).

First we will consider the following problem. Let $\mathbf{P}$ be a probability measure on $(S, \mathcal{S})$, let $\Theta \subset \mathbf{R}^d$ be a $G_\delta$ set such that $\theta_0 \in Int(\Theta)$ and let $g : S \times \Theta \to \mathbf{R}$ be a jointly measurable function. We let $G(\theta) = \mathbf{P}g(\cdot,\theta)$ $G_n(\theta) = \mathbf{P}_n g(\cdot,\theta)$ for $\theta \in \Theta$ where $\mathbf{P}_n$ is the empirical measure based on $X_1, \ldots, X_n$. We make the following assumptions:

(A.1) $G(0) = \sup_{\theta\in\Theta} G(\theta)$.

(A.2) There is a symmetric positive definite bilinear form $A_G$ such that for $\theta$ in a neighborhood of 0

(3.1) $G(\theta) = G(0) - 1/2\, A_G(\theta,\theta) + o(|\theta|^2/\log\log \frac{1}{|\theta|})$.

Without lost of generality we will take $A_G = Id$, i.e. $A_G(\theta,\theta') = \theta\cdot\theta'$.

(A.3) There exists $\Delta : S \to \mathbf{R}^d$ such that $\mathbf{P}|\Delta|^2 < \infty$ and such that, if we define implictly $r(x,\theta)$ by

(3.2) $g(x,\theta) = g(x,0) + \theta\cdot\Delta(x) + |\theta| r(x,\theta)$

then there is $m < \infty$ and functions $r_i(x,\theta)$, $i = 1,\ldots,m$ such that $r(\cdot,\theta) = \sum_{i=1}^m r_i(\cdot,\theta)$ and for some $K > 0$ the classes fo functions $\mathcal{F}_i = \{r_i(\cdot,\theta) : |\theta| \le k\}$ are measurable uniformly bounded VC-graph classes. We let $\mathcal{F} = \{r(\cdot,\theta) : |\theta| \le K\}$.

(A.4) $\mathbf{P}r_i^2(\cdot,\theta) \le [\log\log(\frac{1}{|\theta|})]^{-(2+\delta)}$ for some $\delta > 0$, all $i = 1,\ldots,m$ and all $\theta$ in a neighborhood of 0.

(A.5) There is a r.v. $\theta_n$ that is a estimator of $\theta(=0)$ satisfying

(3.3) $\mathbf{P}_n g(\cdot,\theta_n) + o(n^{-1}) \ge \sup_{\theta\in\Theta}\mathbf{P}_n g(\cdot,\theta)$ where $o(n^{-1})$ is independent of $\omega$ and

(3.4) $\theta_n \to 0$ a.s.

Analogously we have a bootstrap estimator $\hat{\theta}_n^\omega$ satisfying

(3.5) $\hat{\mathbf{P}}_n^\omega g(\cdot,\hat{\theta}_n) + o(n^{-1}) \ge \sup_{\theta\in\Theta}\hat{\mathbf{P}}_n^\omega g(\cdot,\theta)$ and

(3.6) $\hat{\theta}_n^\omega - \theta_n(\omega) \xrightarrow{P} 0$ a.s.

**Remark 3.1.** Note that (A.3) implies that $\mathcal{F}$ is $\mathbf{P}$-Donsker for all $\mathbf{P}$ (e.g. Alexander (1987)). The measurability of $\theta_n$ is guaranted if $(S,\mathcal{S},\mathbf{P})$ is complete by e.g. the section theorem in Cohn (1980), Corollary 8.5.4., page 286). Consider the set

$C = \{(x,\theta) \in S^n \times \Theta : \sum_{i=1}^n g(x_i,\theta) + o(n^{-1}) \ge \sup_{\theta\in\Theta}\sum_{i=1}^n g(x_i,\theta)\}$. Then the section theorem allows as to take a measurable function $\theta_n : S^n \to \Theta$ so that $(x,\theta_n(x)) \in C$ for each $x \in S^n$. Then the estimator in (A.5) is $\theta_n(X_1,\ldots,X_n)$.

**Theorem 3.2.** Under (A.1)-(A.6)

(3.7) $\lim_{n\to\infty}\hat{\mathcal{L}}\ (n^{1/2}(\hat{\theta}_n^{\omega}-\theta_n(\omega))) =_{a.s.} \lim_{n\to\infty}\mathcal{L}\ (n^{1/2}\theta_n)$ which is $N(0, A_G^{-1}(Cov\Delta)A_G^{-1})$.

**Proof.** We proceed in three steps.

**Claim 1.** There exists $c<\infty$ such that, letting $a_n=(n^{-1}\log\log n)^{1/2}$,

(3.8) $\limsup_{n\to\infty}|\theta_n|/a_n \le c$ a.s. and

(3.9) $\hat{\mathbf{P}}\{|\hat{\theta}|_n/a_n\le c\}\to 1$ a.s.

**Proof of claim 1.** Since $\theta_n\to 0$ a.s., (3.1) implies $G(\theta_n)-G(0)<-|\theta_n|^2/4$ for all $n\ge n_0$, for some $n_0<\infty$. Since $\mathbf{P}|\Delta|^2<\infty$ it satisfies the law of the iterated logarithm (LIL) on $\mathbf{R}^d$ . Also by (A.3) the class $\mathcal{F}$ verifies the LIL (Dudley and Phillipp (1983)). Hence there is a constant $k<\infty$ so that

(3.10) $\limsup_{n\to\infty}|(\mathbf{P}_n-\mathbf{P})(\Delta)|/a_n\le k$ a.s. and

(3.11) $\limsup_{n\to\infty}\|\mathbf{P}_n-\mathbf{P}\|_{\mathcal{F}}\ /a_n\le k$ a.s.

These observations, together with (3.1) and the defininition of r give that, for n large,

$0\le\mathbf{P}_n(g(\cdot,\theta_n)-g(\cdot,0))+o(n^{-1})=(\mathbf{P}_n-\mathbf{P})(g(\cdot,\theta_n)-g(\cdot,0))+\mathbf{P}(g(\cdot,\theta_n)-g(\cdot,0))+o(n^{-1})\le$ $(\mathbf{P}_n-\mathbf{P})(\theta_n\cdot\Delta(\cdot)+|\theta_n|r(\cdot,\theta_n))+\mathbf{P}(g(\cdot,\theta_n)-g(\cdot,0))+o(n^{-1})\le 2|\theta_n|a_n(k+o(1))-\theta_n^2/4+o(n^{-1})$ a.s. This inequality implies (3.6) for $c=8k$.

Let $\Delta_n=(\mathbf{P}_n-\mathbf{P})\Delta$ and $\hat{\Delta}_n=(\hat{\mathbf{P}}_n-\mathbf{P}_n)\Delta$. Now, the bootstrap CLT for $\Delta(\cdot)$ gives that for any $c>0$

(3.12) $\hat{\mathbf{P}}\{|\hat{\Delta}|_n/a_n\le c\}\to 1$ a.s.

Similarly, since $\mathcal{F}$ is a.s.-bootstrap **P**-Donsker

(3.13) $\mathbf{P}\{\|\hat{\mathbf{P}}_n-\mathbf{P}_n\|_{\mathcal{F}}\ /a_n\le c\}\to 1$ a.s.

Now proceeding as above, we have, for large n

$$0 \leq \hat{\mathbf{P}}_n(g(\cdot,\hat{\theta}_n)-g(\cdot,0))+o(n^{-1}) \leq |\hat{\theta}_n|[|\hat{\Delta}_n|+|\Delta_n|+\|\hat{\mathbf{P}}_n-\mathbf{P}_n\|_{\mathcal{F}}+\|\mathbf{P}_n-\mathbf{P}\|_{\mathcal{F}}]-|\hat{\theta}_n|^2/4+o(n^{-1})$$

and this, by (3.10)-(3.13) and the LIL for $\mathcal{F}$, gives (3.9).

**Claim 2.** The following limits hold for any $c < \infty$:

(3.14) $(\log\log n)^{1/2} \sup_{|\theta| \leq c\, a_n} |\nu_n(r(\cdot,\theta))| \to 0$ a.s. and

(3.15) $(\log\log n)^{1/2} \sup_{|\hat{\theta}| \leq c\, a_n} |\nu_n(r(\cdot,\theta))| \xrightarrow{P} 0$ a.s.

**Proof of claim 2.** (3.14) is immediate from (A.3), (A. 4) and corollary 2.3 applied to $\mathcal{F}_i$, $i \leq m$. In order to prove (3.15) using Alexander's bound, we must estimate the size of $\alpha_n = \sup\{\mathbf{P}_n(r_i(\cdot,\theta))^2 : |\theta| \leq c\, a_n\}$. For this we use the "square root trick" inequality in Giné and Zinn (1984) lemma 5.2 which gives $\mathbf{P}\{\alpha_n > 4(\log\log n)^{-(2+\delta)}\} \leq (\log\log n)^{\tau} exp(-n/\log\log n)^{1+\delta'})$ for some $\tau > 0$, $\delta' > 0$ and all n large enough. Therefore, eventually $\alpha_n \leq 4(\log\log n)^{-(2+\delta)}$ a.s. this, just as (A.4) in the non-random case, allows us to apply corollary 2.3 conditionally on the sample, and obtain (3.15).

**Claim 3.** $n^{1/2}(\hat{\theta}_n - \Delta_n - \hat{\Delta}_n) \xrightarrow{P} 0$ a.s.

**Proof of claim 3.** By (3.1) and the definition of $\Delta_n$, $\hat{\Delta}_n$, $\theta_n$ and r we have:

$$0 \leq \hat{\mathbf{P}}_n(g(\cdot,\hat{\theta}_n) - g(\cdot,\Delta_n + \hat{\Delta}_n)) + o(n^{-1}) =$$

$$[\hat{\mathbf{P}}_n - \mathbf{P}_n + \mathbf{P}_n - \mathbf{P}](g(\cdot,\hat{\theta}_n) - g(\cdot,\Delta_n + \hat{\Delta}_n)) + \mathbf{P}(g(\cdot,\hat{\theta}_n) - g(\cdot,\Delta_n + \hat{\Delta}_n)) + o(n^{-1}) =$$

$$[\hat{\mathbf{P}}_n - \mathbf{P}_n + \mathbf{P}_n - \mathbf{P}][(\hat{\theta}_n - \Delta_n - \hat{\Delta}_n) \cdot \Delta(\cdot) + |\hat{\theta}_n|\, r(\cdot,\hat{\theta}_n) - |\Delta_n + \hat{\Delta}_n| r(\cdot,\Delta_n + \Delta_n)]$$

$$-|\hat{\theta}_n|^2/2+|\Delta_n+\hat{\Delta}_n|^2/2+o(|\theta_n|^2/\log\log(1/|\hat{\theta}_n|)+o(|\Delta_n+\hat{\Delta}_n|^2/\log\log(1/|\Delta_n+\hat{\Delta}_n|))+o(n^{-1}).$$

So collecting the terms and multipliying by n,

$$(3.16)\quad n|\hat{\theta}_n - \Delta_n - \hat{\Delta}_n|^2/2 \le n^{1/2}|\hat{\theta}_n| n^{1/2}([\hat{\mathbf{P}}_n - \mathbf{P}_n + \mathbf{P}_n - \mathbf{P}](r(\cdot,\hat{\theta}_n))|$$

$$+n^{1/2}|\Delta_n + \hat{\Delta}_n|| n^{1/2}([\hat{\mathbf{P}}_n - \mathbf{P}_n + \mathbf{P}_n - \mathbf{P}](r(\cdot,\Delta_n + \hat{\Delta}_n))| + n\, o(|\hat{\theta}_n|^2/\log\log(1/|\hat{\theta}_n|) +$$

$$n\, o(|\Delta_n + \hat{\Delta}_n|^2/\log\log(1/|\Delta_n + \hat{\Delta}_n|) + o(1).$$

Now (3.9) and (3.15) immediately give $n^{1/2}|\hat{\theta}_n| n^{1/2}([\hat{\mathbf{P}}_n - \mathbf{P}_n + \mathbf{P}_n - \mathbf{P}](r(\cdot,\theta_n))| \xrightarrow{P} 0$ a.s. (3.10), (3.12), (3.14) and (3.15) show

$n^{1/2}|\Delta_n + \hat{\Delta}_n))|| n^{1/2}([\hat{\mathbf{P}}_n - \mathbf{P}_n + \mathbf{P}_n - \mathbf{P}](r(\cdot,\Delta_n + \hat{\Delta}_n))| \xrightarrow{P} 0$ a.s.

and the "o" terms also tend to zero in $\hat{\mathbf{P}}$ a.s. by (3.9), (3.10) and (3.12). Replacing these limits in (3.16) we get

$$(3.17)\quad n^{1/2}(\hat{\theta}_n - \Delta_n - \hat{\Delta}_n) \xrightarrow{P} 0 \text{ a.s.}$$

**Claim 4.** $n^{1/2}(\theta_n - \Delta_n) \to 0$ a.s.

**Proof of claim 4.** The argument leading to (3.17), applied to the inequality $0 \le \mathbf{P}_n(g(\cdot,\theta_n) - g(\cdot,\Delta_n)) + o(n^{-1})$ (which holds by the definition of $\theta_n$, yields

$$n|\theta_n - \Delta_n|^2/2 \le n^{1/2}|\theta_n||\, n^{1/2}([\mathbf{P}_n - \mathbf{P}](r(\cdot,\theta_n))| + n^{1/2}|\Delta_n|\, |[\mathbf{P}_n - \mathbf{P}]r(\cdot,\Delta_n)| +$$

$$n\, o(|\theta_n|^2/\log\log(1/|\theta_n|) + n\, o(|\Delta_n|^2/\log\log(1/|\Delta_n|) + o(1).$$

All the terms at the right side of this inequality tend to zero a.s. by (3.8) and (3.10). Hence, $n^{1/2}(\theta_n - \Delta_n) \to 0$ a.s.

**Proof of Theorem 3.2.** Now (3.17) becomes $n^{1/2}(\hat{\theta}_n - \theta_n - \hat{\Delta}_n) \xrightarrow{P} 0$ a.s. By the bootstrap CLT in $\mathbf{R}^d$ $n^{1/2}\hat{\Delta}_n \xrightarrow{P} N(0, Cov\Delta)$ a.s. And therefore $n^{1/2}(\hat{\theta}_n - \theta_n)$ converges in conditional law a.s. to the same limit. □

Next we give sufficient conditions for the consistency hypothesis (A.5) and (A.6) to hold.

They are slighty stronger than those in Huber (1967). Let, as above, $\mathbf{P} \in \mathcal{P}(S)$, $\Theta \subset \mathbf{R}^d$ be a $G_\delta$ set such that $0 \in Int(\Theta)$, $g : S \times \Theta \to \mathbf{R}$ jointly measurable, and assume, without loss of generality, that $G(0) = 0$ (where $G(\theta) = \mathbf{P}g(\cdot, \theta)$). The conditions that will imply consistency are as follows:

(B.1) $\mathbf{P}|g(\cdot, \theta)| < \infty$ for each $\theta \in \Theta$.

(B.2) There exists $M > 0$ and a nonnegative function $b(\theta)$, $\theta \in \Theta$, such that if

$C = \{\theta \in \Theta : |\theta| \leq M\}$ then

(i) $\inf_{\theta \notin C} b(\theta) = b > 0$.

(ii) $\mathbf{P} \sup_{\theta \notin C} g(\cdot, \theta)/b(\theta) = a < 0$ and

(iii) $\sup_{\theta \in C} |g(\cdot, \theta)| \in L_1(\mathbf{P})$

(B.3) $\sup_{\theta \in C\, |\theta| > \epsilon} G(\theta) < 0$ for all $\epsilon > 0$

(B.4) The class of functions $\mathcal{G} = \{g(\cdot, \theta) : \theta \in C\}$ is a measurable **P**-Glivenko-Cantelli class.

A class of functions $\mathcal{G}$ is **P**-Glivenko-Cantelli if $\|\mathbf{P}_n - \mathbf{P}\|_{\mathcal{G}} \to 0$ a.s. See Giné-Zinn (1984) and Dudley (1984) to conditions of $\mathcal{G}$ to be a measurable **P**-Glivenko-Cantelli class.

**Theorem 3.3.** If **P**, $\mathcal{G}$ and $\theta$ verify conditions (B.1)-(B.4) with $G(0) = 0$, and if $\theta_n$ and $\hat{\theta}_n$ verify the inequalities (3.3) and (3.5) then (3.4) and (3.6) holds.

**Proof.** (3.16) is proved in Huber (1967). To prove (3.17) we first observe that by the bootstrap weak law of large numbers in **R** by (B.2)

(3.18) $\hat{\mathbf{P}}_n \sup_{\theta \notin C} g(\cdot, \theta)/b(\theta) \xrightarrow{P} \mathbf{P} \sup_{\theta \notin C} g(\cdot, \theta)/b(\theta)$ a.s. and

(3.19) $\hat{\mathbf{P}}_n g(\cdot, 0) \xrightarrow{P} \mathbf{P}g(\cdot, 0)$ a.s.

Take $\epsilon > 0$ so small that $(a + \epsilon)b + \epsilon < 0$. Call

$A_n = \{|\hat{\mathbf{P}}_n\{\sup_{\theta \notin C} g(\cdot, \theta)/b(\theta) - \mathbf{P} \sup_{\theta \notin C} g(\cdot, \theta)/b(\theta)| \leq \epsilon\}$ and $B_n = \{|\hat{\mathbf{P}}_n g(\cdot, 0) - \mathbf{P}g(\cdot, 0)| \leq \epsilon\}$

then by (3.16) and (3.17) $\hat{\mathbf{P}}\{A_n\} \to 1$ a.s. and $\hat{\mathbf{P}}\{B_n\} \to 1$ a.s. So $\hat{\mathbf{P}}\{A_n \cap B_n\} \to 1$ a.s.

If $\omega^* \in A_n \cap B_n$ for n large and $\theta \notin C$

$\hat{\mathbf{P}}_n g(\cdot,\theta)) < (a+\epsilon)b(\theta) \leq (a+\epsilon)b < -\epsilon - o(n^{-1}) \leq \hat{\mathbf{P}}_n g(\cdot,0) - o(n^{-1})$.

Hence we get that

(3.20) $\hat{\mathbf{P}}\{\hat{\theta}_n \notin C\} \to 0$ a.s.

Now, since $\mathcal{G}$ is **P**-GC (B.4) and (B.3) (ii) holds, the bootstrap Glivenko -Cantelli theorem in Giné and Zinn (1990a) shows that

(3.21) $|\sup_{\theta\in C,|\theta|>\epsilon} \hat{\mathbf{P}}_n g(\cdot,\theta) - \sup_{\theta\in C,|\theta|>\epsilon} G(\theta)| \leq \sup_{\theta\in C,|\theta|>\epsilon} |(\hat{\mathbf{P}}_n - \mathbf{P})(g(\cdot,\theta)| \xrightarrow{P} 0$ a.s. for all $\epsilon > 0$.

But by an argument, similar to the one done to get (3.18), from (B.3), (3.19), (3.20) and (3.21) we get that $\hat{\mathbf{P}}\{|\theta| > \epsilon\} \to 0$ a.s. □

We now consider the second type of M-estimators. Let $\Theta \subset \mathbf{R}^d$ with $0 \in Int(\Theta)$ $\mathbf{P} \in \mathcal{P}(\mathrm{S})$ and $h : S \times \Theta \to \mathbf{R}^d$ jointly measurable be such that

(C.1) $\mathbf{P}h(\cdot,0) = 0, \text{and} P|H(\cdot,0)|^2 < \infty$.

(C.2) $H(\theta) := \mathbf{P}h(\cdot,\theta)$ is "strongly " differentiable at zero with non-degenerate first derivative. Assuming (without loss of generality) $H'(0) = Id$ the differentiability condition is as follows:

(3.22) $H(\theta) - H(\theta') = \theta - \theta' + o(|\theta - \theta'|)$ as $\theta \to 0$ and $\theta' \to 0$.

(C.3) The classes of functions $\mathcal{F}_i = \{h_i(\cdot,\theta) - h_i(\cdot,0) : |\theta| \leq M\}$ for some $M > 0$, $i = 1,\dots,d$ where $h_i$ denotes the i-th coordinate of h, satisfy the following property: there is $m < \infty$ and uniformly bounded measurable VC-graph classes $g_{ij} = \{g_{ij}(\cdot,\theta) : |\theta| \leq M\}$ such that $h_i(\cdot,\theta) - h_i(\cdot,0) = \sum_{j=1}^m g_{ij}(\cdot,\theta))$.

(C.4) For $i \leq d$, $j \leq m$ $Var_{P} g_{ij}(\cdot,\theta) \leq (\log\log(1/|\theta|))^{-2-\delta}$ for some $\delta > 0$ and $\theta$ in a neighboorhood of 0.

(C.5) There exist symmetric measurable functions $\theta(x_1, \ldots, x_n)$ defined on the support of $\mathbf{P}^n$ such that if $\theta_n := \theta(X_1, \ldots, X_n)$, then

(3.23) $n^{1/2}\mathbf{P}_n h(\cdot, \theta_n) \to 0$ a.s. and $\theta_n \to 0$ a.s.

(C.6) For almost every $\omega \in \Omega$, there exist symmetric random variables $\theta_n^\omega = \theta_n^\omega(X_{n1}^\omega, \ldots, X_{nn}^\omega)$ such that

(3.24) $n^{1/2}\mathbf{P}_n^\omega h(\cdot, \theta_n^\omega) \xrightarrow{P} 0$ a.s and $\hat{\theta}_n^\omega - \theta_n(\omega) \xrightarrow{P} 0$ a.s.

As before, consitency will be handled separately.

**Theorem 3.4.** Under (C.1)-(C.6),

(3.25) $\lim_{n\to\infty} \hat{\mathcal{L}}\ (n^{1/2}(\hat{\theta}_n - \theta_n)) =_{a.s.} \lim\ _{n\to\infty} \mathcal{L}\ (n^{1/2}(\theta_n))$ which is $N(0, cov_P h(\cdot, 0))$.

**Proof.** First we show that

(3.26) $n^{1/2}\mathbf{P}h(\cdot, \theta_n) + n^{1/2}\mathbf{P}_n h(\cdot, 0) \to 0$ a.s.

since $H'(0)$ is not degenerate and $\theta_n \to 0$ a.s. ((3.24) and (3.25)) there exist $c > 0$ and $n_0(\omega) < \infty$ a.s. such that for $n \geq n_0$ then $c|\theta_n| \leq |H(\theta_n)|$. Hence, using (3.21) and the LIL for random vectors in $\mathbf{R}^d$ and for empirical processes we have

$n^{1/2}c|\theta_n| \leq n^{1/2}|H(\theta_n)| \leq |n^{1/2}(\mathbf{P}_n - \mathbf{P})h(\cdot, \theta_n)| + n^{1/2}|\mathbf{P}_n h(\cdot, \theta_n)|$

$\leq |n^{1/2}(\mathbf{P}_n - \mathbf{P})h(\cdot, 0)| + |n^{1/2}(\mathbf{P}_n - \mathbf{P})(h(\cdot, \theta_n) - h(\cdot, 0))| + n^{1/2}|\mathbf{P}_n h(\cdot, \theta_n)| = O((\log\log n)^{1/2})$ a.s.

Hence, for some $c < \infty$

(3.27) $\limsup\ _{n\to\infty} n^{1/2}|\theta_n|/(\log\log n)^{1/2} \leq c$ a.s.

Theorem 2.8 gives, as in the proof of Theorem 3.2., that for all $c > 0$,

(3.28) $\sup\ _{|\theta| \leq \epsilon\alpha_n} n^{1/2}|(\mathbf{P}_n - \mathbf{P})(h(\cdot, \theta) - h(\cdot, 0))| \to 0$ a.s.

which, by (3.26), implies $n^{1/2}(\mathbf{P}_n - \mathbf{P})(h(\cdot, \theta_n) - h(\cdot, 0))| \to 0$ a.s.

Hence, by (C.1) and (3.23), $n^{1/2}(\mathbf{P}h(\cdot,\theta_n)+\mathbf{P}_nh(\cdot,0)) \to 0$ a.s. which is claim (3.26).

By the bootstrap CLT for empirical processes (Theorem 2.1) and (C..3), for every sequence $b_n \to \infty$ $\hat{\mathbf{P}}\{n^{1/2}|(\hat{\mathbf{P}}_n - \mathbf{P}_n(H(\cdot,\hat\theta_n) - h(\cdot,0))| \le b_n\} \to 1$ a.s.

We can use this and (3.24) in a decomposition similar to (3.25). And obtain that there is $c < \infty$ such that

(3.29) $\hat{\mathbf{P}}\{n^{1/2}|\hat\theta_n|/(\log\log n)^{1/2} \le c\} \to 1$ a.s.

We consider now, in analogy with (3.26),

$$n^{1/2}(\mathbf{P}h(\cdot,\hat\theta_n)+\hat{\mathbf{P}}h(\cdot,0)) = n^{1/2}\hat{\mathbf{P}}_nh(\cdot,\hat\theta_n) - n^{1/2}((\hat{\mathbf{P}}_n-\mathbf{P}_n)+\mathbf{P}_n-\mathbf{P}))(h(\cdot,\hat\theta_n)-h(\cdot,0)).$$

The first term at the right hand side tends to zero in $\hat{\mathbf{P}}$, a.s . by (3.24). The next term, $\hat\nu_n(h(\cdot,\hat\theta_n)-h(\cdot,0))$, also converges to zero in $\hat{\mathbf{P}}$,a.s by Theorem 2.1. (the bootstrap CLT implies, by (C.4),

$\lim_{\delta\to 0} \limsup_{n\to\infty}\hat{\mathbf{P}}\{ \sup_{|\theta|<\delta}\hat\nu_n(h(\cdot,\theta)-h(\cdot,0))| > \epsilon\} = 0$ a.s.for all $\epsilon > 0$ and $\hat\theta_n \xrightarrow{P} 0$ a.s. by (3.23) and (3.24). And the same is true for the last term $\nu_n(H(\cdot,\hat\theta_n)-H(\cdot,0)$ by (3.29) and (3.30). therefore $n^{1/2}\mathbf{P}h(\cdot,\hat\theta_n)+n^{1/2}\hat{\mathbf{P}}_nh(\cdot,0) \xrightarrow{P} 0$ a.s this limit and (3.24) give, by substration,

(3.30) $n^{1/2}(H(\hat\theta_n)-H(\theta_n)) - n^{1/2}(\hat{\mathbf{P}}_n-\mathbf{P}_n)h(\cdot,0) \xrightarrow{P} 0$ a.s.

This shows that for almost every $\omega$ the sequence $\{n^{1/2}(H(\hat\theta_n^\omega)-H(\theta_n(\omega))\}$ in $\hat{\mathbf{P}}$-stochastically bounded: it converges weakly by bootstrap CLT in $\mathbf{R}^d$. since $\hat\theta_n - \theta_n \xrightarrow{P} 0$ a.s. ((3.24)) hypothesis (C.2) implies $\hat{\mathbf{P}}\{|\hat\theta_n-\theta_n| > 2|H(\hat\theta_n)-H(\theta_n\} \to 0$ a.s. Hence the sequence $\{n^{1/2}|\hat\theta_n^\omega - \theta_n(\omega)|\}$ is $\hat{\mathbf{P}}$ stochatiscally bounded,$\omega$ a.s. Then, $n^{1/2}o(|\hat\theta_n-\theta_n|) \xrightarrow{P} 0$ a.s and (3.31) and (3.22) give

$n^{1/2}(\hat\theta_n-\theta_n)+n^{1/2}(\hat{\mathbf{P}}_n-\mathbf{P}_n)h(\cdot,0) \xrightarrow{P} 0$ a.s. Now (3.25) follows by the bootstrap CLT in $\mathbf{R}^d$ □

The consistency conditions of Huber (1967, case B) not only give consistency of $\theta_n$, but also consistency of $\hat{\theta}_n$. The proof is similar to that of theorem 3.2.

**Theorem 3.5.** If h is jointly measurable and if $\mathbf{P}_n h(\cdot,\theta_n) \to 0$ a.s. $\hat{\mathbf{P}}_n h(\cdot,\hat{\theta}_n) \xrightarrow{P} 0$ a.s. for random variables $\theta_n$ , $\hat{\theta}_n$, then $\theta_n \to 0$ a.s. and $\theta_n - \hat{\theta}_n \xrightarrow{P} 0$ a.s., assuming the following conditions hold:

(1) h is **P**-a.s. continuous in $\theta$.

(2) $H(\theta)$ exists for all $\theta \in \Theta$ and has a unique zero at $\theta = 0$.

(3) There exists a continuous function b on $\Theta$, bounded away from zero, and $C = \{\theta : |\theta| \leq M\} \cap \Theta$ for some $M < \infty$, such that $\mathbf{P} \sup_{\theta \in \theta} |h(\cdot,\theta)|/b(\theta) < \infty$, $\inf_{\theta \notin C} |H(\theta)|/b(\theta) \geq 1$ and $\mathbf{P} \sup_{\theta \notin C} (h(\cdot,\theta) - h(\theta))/b(\theta) < 1$.

(4) $\mathbf{P}\sup_{\theta \in C} |h(\cdot,\theta)| < \infty$ and $\{h(\cdot,\theta) : \theta \in C\}$ is a **P**-Glivenko -Cantelli class of functions.

Theorem 3.4 applies to monotone functions h with a significant symplifycation of the hypotheses:

**Theorem 3.6.** Let $h : \mathbf{R} \to \mathbf{R}$ be a bounded non-increasing function and let **P** be a probability measure on **R**. We let $h(x,\theta) = h(x - \theta)$ $x, \theta \in \mathbf{R}$, assume

(D.1) If $H(\theta) = \mathbf{P}h(\cdot,\theta)$, then $H(0) = 0$, $H'(0) > 0$ and

(3.31) $\lim_{r,s \to 0} [H(r) - H(s)]/(r - s) = H'(0)$

(D.2) There is a $\delta > 0$ such that for $|\theta| \leq \delta$ $Var_P(h(\cdot,\theta) - h(\cdot,0)) \leq [\log\log(1/|\theta|)]^{-(2+\delta)}$

(D.3) If C is the set of discontinuity points of $h(\cdot,0)$, $C_\delta = \{x : \exists y \in C \ s.t. \ |x - y| < \delta\}$ then **P** is continuous on $C_\delta$ for some $\delta > 0$. Then $\lim \mathcal{L}\ (n^{1/2}(\theta_n - \theta_n)) =_{a.s.} \lim_{n \to \infty} \mathcal{L}\ (n^{1/2}\theta_n) = N(0, Var_P(h/H'(0))$ where $\theta_n = \inf\{\theta : \mathbf{P}_n h(\cdot,\theta) > 0\}$ and $\hat{\theta}_n = \inf\{\theta : \hat{\mathbf{P}}_n h(\cdot,\theta) > 0\}$ .

**Proof.** (C.1) and (C.2) are (D.1). (C.3) holds because $\{h(\cdot,\theta) : \theta \in \mathbf{R}\}$ is a VC-graph class. $\{ \{(x,t) : 0 \le t \le h(x,\theta)\} : \theta \in \mathbf{R}\}$ and $\{ \{(x,t) : h(x,\theta) \le t \le 0\} : \theta \in \mathbf{R}\}$ are each one a class of sets ordered by inclussion. So they are VC classes. Since the union of VC classes is VC, $\{h(\cdot,\theta : \theta \in \mathbf{R}\}$ is a VC-graph class. (C.4) is (D.2).

As far as (C.5) and (C.6), we first prove that

(3.32) $\theta_n \rightarrow 0$ a.s. and $\theta_n - \hat{\theta}_n \xrightarrow{P} 0$ a.s.

Given $\epsilon > 0$. By (D.1) $H(-\epsilon) < 0 < H(\epsilon)$. Since $H(\cdot,\pm\epsilon) \in L_1(\mathbf{P})$, by the LLN and the bootstrap LLN $\mathbf{P}_n h(\cdot,\pm\epsilon) \rightarrow H(\pm\epsilon)$ a.s. $\hat{\mathbf{P}}_n h(\cdot,\pm\epsilon) \xrightarrow{P} H(\pm\epsilon)$ a.s. So (3.31) follows.

Let $|\theta| \le \delta/2$. The sample points $X_i$ satisfying $X_i - \theta \in C_{\delta/2}$ are all a.s. different by continuity of $\mathbf{P}$ on $C_\delta$. Moreover $H(x)$ is continuous at $x = X_i - \theta$ if $X_i - \theta \notin C_{\delta/2}$. Note $H_n(\theta+) - H_n(\theta-) = \sum_{i:X_i-\theta\in C_{\delta/2}} h(X_i + \theta-) - h(X_i - \theta+) \le Variation(h) \le 2\|h\|_\infty$. Hence the function $\mathbf{P}_n h(\cdot,\theta)$ has a jump at $\theta$ of size at most $2\|h\|_\infty n^{-1}$. This proves by (3.33) that $n^{1/2}\mathbf{P}_n h(\cdot,\theta_n) \rightarrow 0$ a.s. and $n^{1/2}\hat{\mathbf{P}}_n h(\cdot,\hat{\theta}_n) \xrightarrow{P} 0$ a.s. □

**Remark 3.7.** We refer to Romo (1990) for the study of the bootstrap in probability of M-estimators. I thank him for some conversations about his work that induced me to study this topic.

## 4. Examples.

**Example 4.1.** (Spatial medians). The spatial median of $\mathbf{P} \in \mathcal{P}\ (\mathbf{R}^d)$ is the value $\theta(\mathbf{P})$ of $\theta$ that minimizes $\mathbf{P}|x - \theta|$ where $|x|$ means the Euclidean distance in $\mathbf{R}^d$. So this estimator are in the scheme of theorems 3.2 and 3.5 taking $g(x,\theta) = -|x - \theta|$. If d=1, the spatial median is the ususal median. The set of median of $\mathbf{P}$, which is convex consists of a single point unless $\mathbf{P}$ is concentrated on a line (and has more than one median there). Note that if $\theta$ and $\theta'$ are two different spatial medians then for any $0 < \alpha < 1$ $|\alpha\theta + (1-\alpha)\theta' - x| \leq \alpha|\theta - x| + (1-\alpha)|\theta' - x|$. Thus $\mathbf{P}|\alpha\theta+(1-\alpha)\theta'-x| = \mathbf{P}|\theta-x|$. Hence $|\alpha\theta+(1-\alpha)\theta-x| = \alpha|\theta-x|+(1-\alpha)|\theta-x|$ a.s. But this implies that almost every x is in the part of the line through $\theta$ and $\theta'$ exterior to the segment $\theta$ and $\theta'$. We assume that $\mathbf{P}$ is not concentrated on a line and has a bounded density in a neighborhood of $\theta(\mathbf{P})$. From the fact that for $x \neq 0$ $| \, |x| - |x-\theta| - (\theta \cdot x)|x|^{-1} + 2^{-1}|\theta|^2|x|^{-1} - 2^{-1}|\theta \cdot x|^2|x|^{-3}| \leq 2|\theta|^3|x|^{-2} \wedge 3|\theta|^2|x|^{-1}$

it follows changing to spherical coordinates that

$G(\theta) = G(0) - \mathbf{P}[2^{-1}|\theta|^2|x|^{-1} - 2^{-1}|\theta \cdot x|^2|x|^{-3}] + O(|\theta|^3)$ if $d \geq 3$

$G(\theta) = G(0) - \mathbf{P}[2^{-1}|\theta|^2|x|^{-1} - 2^{-1}|\theta \cdot x|^2|x|^{-3}] + O(|\theta|^3 \log(1/|\theta|)$ if $d = 2$

Hence (A.2) holds. Pollard (1984, page 153) shows that $\mathcal{F} = \{r(x,\theta) = [|x| - |x-\theta| - (\theta \cdot x)/|x|]/|\theta|) : |\theta| \neq 0\} \cup \{r(x,0) \equiv 0\}$ is the sum of two VC graph classes, so that (A.3) is satisfied. Moreover, it is easy to see that $|r(x,\theta)| \leq 2 \wedge (2|\theta|/|x|)$ hence if $\mathbf{P}$ has a density f on $\{|x| < \delta\}$ bounded by M, we have $\mathbf{E}r^2(x,\theta) \leq k(d,M)|\theta|^2$ for $\theta$ near to 0 where $k(d,M)$ is a constant depending on d and M. So, (A.4) holds too. finally, taking $b(\theta) = |\theta|$ conditions (B.1)-(B.3) are satisfied, and it is easy to see (as in Pollard (1984), page 153) that $\{|x| - |x-\theta| : |\theta| \leq M\}$ is a bounded VC-graph class, hence a Glivenko -Cantelli class. The conclusion from Theorem 3.1. is then that the central limit theorem for the spatial median can be bootstrapped a.s. assuming $\mathbf{P}$ is not concentrated on a line and $\mathbf{P}$ has a bounded density in a neighborhood of $\theta(\mathbf{P})$.

**Example 4.2.** (k-means in $\mathbf{R}$). Given $\mathbf{P}$ a k-mean of $\mathbf{P}$, $1 \leq k < \infty$, is an ordered set of k points in $\mathbf{R}$, $\theta_1 \leq \cdots \leq \theta_k$, that minimizes the function $\mathbf{P} \min_{i \leq k}(x - \theta_i)^2$ ( Hartigan (1978), Pollard (1981, 1982) Cuesta and Matrán (1988)). This notion has also a meaning in $\mathbf{R}^d$ and even in infinite dimensions. Since our purpose is only to illustrate the use of the previous theorems, we restrict our discussion to the simpler one-dimensional case. We make the following assumptions on **P**:

(1) $\mathbf{P}\mathbf{X}^2 < \infty$.

(2) **P** has a unique k-median $\mu$ and it consists of k distinct points $(\mu_1, \ldots, \mu_k)$ with $\mu_1 < \cdots < \mu_k$.

(3) **P** has differentiable density at the points $(\mu_i + \mu_{i+1})/2$, $i = 1, \ldots, k-1$

We let $\Theta = \{(\theta_1, \ldots, \theta_k) \in \mathbf{R}^k : \theta_1 \leq \cdots \leq \theta_k\} \quad g(x, \theta) = \min_{i \leq k}(x - \theta_i)^2$. By a compactness argument there always exists, for each $(x_1, \ldots, x_n) \in \mathbf{R}^n$ a point in $\Theta$ that minimizes $n^{-1}\sum_{i=1}^{n} \delta_{x_i} g(\cdot, \theta)$. By the section theorem we can take a $\theta(x_1, \ldots, x_n)$ measurable.

Pollard (1982) proved a CLT for $\theta_n - \mu$. We will show here that Theorem 3.2 implies that Pollard's CLT can be bootstrapped a.s. under conditions (1)-(3). For consistency we follow Cuesta and Matrán (1988). They show that if $\mathbf{Z}_n, \mathbf{Z}_0$ are B-valued random variables, B a uniformly convex Banach space, such that $\mathbf{Z}_n \rightarrow Z_0$ a.s. $Z_0$ has a unique k-mean $:= \theta(\mathcal{L}\ (Z_0))$ and $\mathbf{E} \min_{i \leq k} ||\mathbf{Z}_n - \mu_i||^2 \rightarrow \mathbf{E} \min_{i \leq k} ||\mathbf{Z}_0 - \mu_i||^2 < \infty$, then $\theta(\mathcal{L}\ (\mathbf{Z}_n)) \rightarrow \theta(\mathcal{L}\ (Z_0))$. They apply this result and a Skorohod representation to show consistency of $\theta_n$ under hypotheses (1) and (2). Using this argument conditionally on $\omega$ we can can get the bootstrap of the consistency of k-means because if $\mathcal{C} = \{(-\infty, t] : t \in \mathbf{R}\}$ then by the bootstrap law of large numbers $||\mathbf{P}_n^{\omega} - \mathbf{P}||_{\mathcal{C}} \xrightarrow{P} 0$ a.s. and $\hat{\mathbf{P}}_n^{\omega} \min_{i \leq k}(x - \mu_i)^2 \xrightarrow{P} \mathbf{P} \min_{i \leq k}(x - \mu_i)^2$ a.s.

It is easily checked that if **P** has differentiable density f at $(\mu_i + \mu_{i+1})/2$ , $i = 1, \ldots, k-1$, then $G(\theta)$ is three times differentiable at $\theta = \mu$, hence condition (A.2) holds. If we let $M_1 = (-\infty, (\mu_1 + \mu_2)/2]$ $M_j + ((\mu_j + \mu_{j-1})/2, (\mu_j + \mu_{j+1})/2$ $j = 2, \ldots, k-1$ and $M_k((\mu_{k-1} + \mu_k)/2, \infty)$, and if we define $A_j$ for $j = 1, \ldots, k$ in the same way but replacing $\mu$ by $\theta$ we then have:

$\Delta(x) = (-2(x - \mu_1)I_{M_1}(x), \ldots, -2(x - \mu_k)I_{M_k}(x))$ and

$r(x, \theta) = \sum_{i,j}(I_{M_i A_j}(x))(2(\theta_i - \theta_j)x + \mu_i^2 - 2\mu_i\theta_i + \theta_j^2)/|\theta - \mu|$ (Pollard(1982)). Using the obvious facts that:

(a) if $|\mu-\theta| \le k^{-1/2}\min(\mu_{i+1}-\mu_i)/2$ then $M_iA_j=\emptyset$ unless $|i-j|\le 1$, and

(b) $I_{M_iA_{i+1}} \subset [(\mu_i+\mu_{i+1})/2-|\theta-\mu|,(\mu_i+\mu_{i+1})/2]$

$I_{M_iA_{i-1}} \subset [(\mu_i+\mu_{i+1})/2,(\mu_i+\mu_{i+1})/2+|\theta-\mu|]$, we obtain

$|r(x,\theta)| \le |\mu-\theta| + 4\sum_{i=1}^{k-1}(I_{M_iA_{i+1}}+I_{M_{i+1}A_i})|\theta_i-\theta_{i+1}|$

This inequality implies that $\sup_{|\theta|\le K}\|r(x,\theta)\|_\infty < \infty$ for all $k<\infty$ and that $\mathbf{P}r^2(\cdot,\theta) \le c|\theta-\mu|$ for all $\theta$ in a neighborhood f $\mu$ and for some $c<\infty$. Hence (A.4) holds. For each $\theta$, $r(x,\theta)$ is the sum of $k^2$ or less functions that are linear on an interval and zero outside it . Hence condition (A.3) is also satisfied. Therefore Theorem 3.2 applies and it follows that, under conditiones (1)-(3) on **P**, the CLT for k-means of Pollard (1982) bootstrap a.s. ( for d=1).

**Example 4.3** (A kind of Winsorized mean). Huber (1964) presents the M-estimator defined as in the way of the theorem 3.8 with $h(x) = -k\ I_{x\le -k} + x\ I_{-k<x\le k} + k\ I_{k<x}$. Note that $H(\theta)$ is continuous and nondecreasing with $\lim_{\theta\to-\infty} H(\theta)=0$ $\lim_{\theta\to\infty} H(\theta)=1$. So there is a $\theta_0$ so that $H(\theta_0)=0$ Assume that

(1) **F** is continuous at $\theta_0+k$ and $\theta_0-k$

(2) $\mathbf{F}(\theta_0+k)-\mathbf{F}(\theta_0-k)\ne 0$

Then under this conditions $n^{1/2}\theta_n \longrightarrow N(0,\sigma^2)$. Without additional conditions we get the bootstrap a.s. from the theorem 3.8. A calculation shows $H(\theta)=k-\int_{-k+\theta_0}^{k+\theta_0}\mathbf{F}(x)dx$. By the condition (1) $H(\cdot)$ satisfies (3.30) with derivative $\mathbf{F}(\theta_0+k)-\mathbf{F}(\theta_0-k)$. Since $|h(x,\theta)-h(x,\theta')| \le |\theta-\theta'|$, (D.2) holds.

# REFERENCES

DE ACOSTA, A. and GINÉ, E. (1979). Convergence of moments and related functionals in the central limit theorem in Banach spaces. *Zeits. Wahrs. Verw. Geb.* 48 213-231.

ALEXANDER, K. S. (1984) . Probability inequalities for empirical processes and a law of the iterated logarithm. *Ann. Probab.* **12** 1041-1067.

ALEXANDER, K. S. (1987). The central limit theorem for empirical processes on Vapnik-Cervonenkis classes. *Ann. Probab.* **15** 178-203.

ARAUJO, A. and GINÉ, E. (1980). *The central limit theorem for real and Banach valued random variables.* Wiley, New York.

ARCONES, M. A. and GINÉ, E. (1989a). The bootstrap of the mean with arbitrary bootstrap sample size. *Ann. Inst. Henri Poincaré* 25 457-481.

ARCONES, M. A. and GINÉ, E. (1989b). Some bootstraps test of symmetry for univariate simmetric distributions. To appear in *Ann. of Statistics.*

ARCONES, M. A. and GINÉ, E. (1990a). The bootstrap central limit theorem for U and V statistics. To appear in *Ann. of Statistics.*

ARCONES, M. A. and GINÉ, E. (1990b). On the bootstrap of M-estimators and other statistical functionals. To appear in the Proceeding of the IMS Bootstrap Conference, East Lansing 1990.

ARCONES, M. A. and GINÉ, E. (1991). Additions and correction to "The bootstrap of the mean with arbitrary bootstrap sample size". To appear in *Ann. Inst. Henri Poincaré.*

ATHREYA, K. B. (1984). Bootstrap of the mean in the infinite variance case I.Technical Report 86-22. *Dep. of Statist. Iowa State Univ.*

ATHREYA, K. B. (1985). Bootstrap of the mean in the infinite variance case II.Technical Report **86-21**. *Dep. of Statist. Iowa State Univ.*

ATHREYA, K. B. (1987). Bootstrap of the mean in the infinite variance case. *Ann Statist.* **15** 724-731.

BABU, G. J. (1984). Bootstapping statistics with linear combinations of chi-squares as weak limit. *Sankhya* **46** 86-93.

BERAN, R. (1984). Bootstrap method in statistics. Jber. d. Dt. Math. Verein. **86**, 14-30.

BERK, R.H. (1966). Limiting behavior of posterior distributions where the model is incorrect. *Ann. Math. Statis.* **37** 51-58.

BICKEL, P. J. and FREEDMAN, D. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9** 1196-1217.

BICKEL, P. J. and HODGES, J. L. Jr. (1967). The asymptotic theory of Galton's test and a related example estimate of location. *Ann. Math. Statist.* **38** 73-89.

BILLINGSLEY, P. (1969). *Convergence of probability measures.* Wiley, New York.

BORELL, C. (1974). Convex measures on locally convex spaces. *Ark. Mat.* **12** 239-253.

BRETAGNOLLE, J. (1983). Lois limites du bootstrap de certaines fonctionnelles. *Ann. Inst. H. Poincaré* **3** 281-296.

CHEN, J. and RUBIN, H. (1984). A note on the behavior of sample statistics when the population mean is infinite. *Ann. Prob.* **12** 256-261.

COHN, D. L. (1980). *Measure theory.* Birkhauser, Boston.

CSÖRGÖ, S. and MASON, D. (1989). Bootstrapping empirical functions, *Ann. Statis.* **17** 1447-1471.

DUDLEY, R. M. (1984). A course on empirical processes. *Lect. Notes in Math.* **1097** 1-142. Springer , New York.

DUDLEY, R. M. (1985). An extended Wichura theorem, definitions of Donsker class and weighted empirical distributions. *Lect. Notes in Math.* **1153** 141-178. Springer, New York.

DUDLEY, R. M. and PHILIPP, W. (1983). Invariances principles for sums of Banach space valued random elemnts and empirical processes. *Z. Wahrsch. v. Geb.* **62** 509-552.

EFRON, B. (1979). Bootstrap Methods: Another Look at the Jacknife. *Ann. Statist.* **7** 1-26.

FELLER, W. (1971). *An Introduction to Probability Theory and its applications.* II, Wiley, New York.

FILIPPOVA, A.A. (1961). Mises' theorem on the asymptotic behavior of functional of empirical distribution functions and its statistical applications. *Theor. Probability Appl.* **7** 24-57.

FINE, T. (1966). On the Hodges and Lehmann shift estimator in the two sample problem. *Ann. Math. Statist.* **37** 1814-1818.

GINÉ, E. and ZINN, J. (1984). Some limit theorems for empirical processes. *Ann. Prob.* **12** 929-989.

GINÉ, E. and ZINN, J. (1989a). Neccesary conditions for the bootstrap of the mean, *Ann. Statist.* **17** 684-691.

GINÉ, E. and ZINN, J. (1989b). Gaussian characterization of uniform Donsker classes of functions. To appear in *Ann. Statist.*

GINÉ, E. and ZINN, J. (1990a). Bootstrap of empirical measures. *Ann. Prob.* **18** 851-869.

GINÉ, E. and ZINN, J. (1990b). Marcinkiewicz type laws of large numbers for U-statistics. To appear.

HOEFFDING, W. (1948a). A class of statistics with asymptotically normal distribution . *Ann. Math. Statist.* **19** 293-325.

HOEFFDING, W. (1948b). A non-parametric test of independence. *Ann. Math. Statist.* **19**

546-557.

HODGES, J.L., Jr. (1965). Efficency in normal samples and tolerance of extreme values for some estimates of location. *Proc. Fith Berkely Sym Prob. Statist.* **1** 163-186.

HODGES, J.L., Jr.and LEHMAN, E. (1963). Estimates of location based on ranks tests. *Ann. Math. Statist.* **34** 598-611.

HOFFMANN-JORGENSEN, J. (1974) Sums of independent Banach space valued random variables.

HOFFMANN-JORGENSEN, J., SHEPP, L.A. and DUDLEY, R. M. (1979). On the lower tail of Gaussian seminorms. *Ann. Prob.* **7** 319-342.

HUBER, P. J. (1964). Robust estimatation of a location parameter. *Ann. Math. Stat.* **35** 73-101.

HUBER, P. J. (1967). The behaviour of maximum likelihood estimates under non-standard condition. *Proc. Fifh Berkeley Symposium on Mathematical Statistics and Probability* **1** 221-233. Univ. of California Press, Berkeley.

LOÈVE, M. (1963). *Probability Theory*, 3rd Edition, VanNostrand-Reinhold, Princeton, New Jersey.

MASON, D.M.; SHORAK, G. R. and WELLNER, J. A. (1983). Strong limit theorems for oscillation moduli of the uniform empirical process. *Z. Wahrs. verw. Gebiete* **65** 83-97.

VON MISES, R. (1947). On the asymptotic distribution of differentiable statistical functions. *Ann. Math. Statist.* **18** 309-348.

NEVEU, J. (1968). *Processus aléatoires gaussianes.* Presses de l'université de Montréal.

POLLARD, D. (1981). Strong consistency of k-means clustering. *Ann. Statist.* **9** 135-140.

POLLARD, D. (1982). A central limit theorem for k-means clustering. *Ann. Probability.* **10** 919-926.

POLLARD, D. (1985). New ways to prove central limit theorems. *Econometric Theory* **1** 295-314.

RAO, P. V., SCHUSTER, E. F. and LITTELL, R. C. (1975). Estimation of shift and center of symmetry based on Kolmorov-Smirnov Statistics. *Ann. Statis.* **3** 862-873.

REEDS, J. A. (1976). On the definition of von Mises functional. *Reserch Report S-44.* Dept Statist., Harvard Univ.

ROMANO, J. H. (1988). A bootstrap revival of some parametric distance tests. *J. Amer. Statist. Assoc.* **83** 698-708.

ROMO, J. A. (1990). Manuscript.

RUBIN, H. and VITALE, R. A. (1980). Asymptotic distribution of symmetric statistics. *Ann. Statist.* **8** 165-170.

SCHUSTER, E. F. (1975) . Estimating the distribution function of asymmetric distribution. *Biometr.*62 631-635.

SCHUSTER, E. F. (1987) . Identifying the closest symmetric distribution of r density function. *Ann. Statis.* **15** 865-874.

SCHUSTER, E. F. and BARKER, E. F. (1973). A new nonparametric estimator of the center of a symmetric distribution. *Ann. Statis.* **1** 1096-1104.

SCHUSTER, E. F. and BARKER, R. C. (1987). Using bootstrap in testing symmetry versus asymmetry. *Communi. Sta. Simul.* **16** 69-84.

SERFLING, R. J. (1980). *Approximation theorems of mathematical statistics.* Wiley, New York.

SHORACK, G. R. (1972). Functions of order statistics. *Ann. Math. Statist.* **43** 412-427.

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical processes with applications to statistics.* Wiley, New York.

STOUT, J. A. (1974). *Almost Sure Convergence.* Academic Press, New York.

STUTE, J. A. (1982). Oscillation behavior of empirical processes. *Ann. Probab.* **10** 86-107.

WU, W., CARLSTEIN, E. and CAMBANIS, S. (1989). Bootstrapping the sample mean for data wit infinite variance. Prepint.

# VITA

Miguel Angel Arcones was born in Segovia (Spain), March 1, 1963. He obtained his "Licenciatura" in Mathematics at the Universidad Autónoma de Madrid in June, 1986. He spent the following year at the Universidad Complutense de Madrid as a graduate student. He was a graduate student and a teaching assistant at Texas A & M University during the course 1987-88. He has been a graduate student at the Department of Mathematics at CUNY since september 1988. During the year 1988-89 he was awarded a scholarship from the US-Spain Joint Committee. He was a lecturer at the College of Staten Island during the academic year 1989-90. Presently he is a lecturer at the University of Connecticut. He is the coauthor of nine research articles.