INTRODUCTION: $\psi$-STATISTICS OF A TRANSFORMATION
HOEFFDING’S AND M-C DECOMPOSITIONS
FCLT

FUNCTIONAL LIMIT THEOREMS FOR VON MISES STATISTICS OF A MEASURE PRESERVING TRANSFORMATION

Mikhail Gordin

V.A. Steklov Institute of Mathematics
Saint Petersburg

Limit theorems
for dependent data
and applications

Paris, June 21 – 23, 2010
Outline

1. INTRODUCTION: V-STATISTICS OF A TRANSFORMATION
2. HOEFFDING’S AND M-C DECOMPOSITIONS
3. FCLT
INTRODUCTION: V-STATISTICS OF A MEASURE PRESERVING TRANSFORMATION

(after a joint work with Herold Dehling and Manfred Denker)
V-statistics

Let $T$ be a measure preserving transformation of a probability space $(\Omega, \mathcal{F}, P)$. Choose a point $\omega \in \Omega$ and consider its $n$–orbit

$$\omega, T\omega, \ldots, T^{n-1}\omega.$$ 

From statistician’s point of view this is a sample of size $n$. Let us consider, for a certain measurable symmetric function $h : \Omega^d \to \mathbb{R}$, the expression

$$\sum_{1 \leq i_1 < n, \ldots, 1 \leq i_d \leq n} h(T^{i_1}\omega, \ldots, T^{i_d}\omega).$$ (1)

Such a functional will be called a $V$–statistic (or von Mises statistic) of degree $d$ with the kernel $h$. 

Mikhail Gordin

FUNCTIONAL LIMIT THEOREMS
Let $X = (X_n)_{n \in \mathbb{Z}}$ be a strictly stationary real-valued sequence. Every such $X$ admits a representation of the form

$$X_n = f \circ T^n, \; n \in \mathbb{Z},$$

where $T$ is a measure preserving invertible transformation of a certain probability space and $f$ is a measurable function. Let $H : \mathbb{R}^d \rightarrow \mathbb{R}$ be a Borel measurable function. If we set

$$h(\omega_1, \ldots, \omega_d) = H(f(\omega_1), \ldots, f(\omega_d)),$$

we arrive from (1) at the standard expressions for a $V$-statistic:

$$\sum_{1 \leq i_1 < n, \ldots, 1 \leq i_d \leq n} H(X_{i_1}, \ldots, X_{i_d}). \quad (2)$$
Dynamics can be used as follows to generate the function

$$\omega \mapsto h(T_1^{i_1}\omega, \ldots, T_d^{i_d}\omega).$$

First, we consider an action of $d$ commuting copies $T_1, \ldots, T_d$ of $T$ on some set $Y \subset \Omega^d$ to produce terms of the form

$$h(T_1^{i_1}\omega_1, \ldots, T_d^{i_d}\omega_d).$$

Second, we restrict the constructed function to the principal diagonal $D = \{(\omega, \ldots, \omega) : \omega \in \Omega\} \subset \Omega^d$ and obtain the desired term. The requirements which $Y$ must satisfy are:

i) $T_k Y \subset Y, k = 1, \ldots, d$;  
ii) $D \subset Y$.

We choose as $Y$ the entire space $\Omega^d$ with the product measure $P^d$ and the componentwise action of copies of $T$. 

Mikhail Gordin
Let \( h : \Omega^d \to \mathbb{R} \) be (an equivalence class) of a certain measurable function on \( \Omega^d \). Consider the set

\[
\bigcup_{(n_1, \ldots, n_d) \in \mathbb{Z}^d} \left\{ (T_1^{n_1}\omega, \ldots, T_d^{n_d}\omega), \omega \in \Omega \right\}
\]

doing measure zero. For \( d = 2 \) this is the graph of the orbital equivalence relation of \( T \).

What is the correct restriction of \( h \) to subsets of this set?

In general, no idea.

However, the restriction problem is easily solvable for kernels \( h \) which are products of functions in one variable, or can be nicely approximated by sums of such functions.

We will use such an approximation.
Seminal papers
Hoeffding (1948): U-statistics for i.i.d. variables; Hoeffding’s decomposition
von Mises (1949): V-statistics for i.i.d. variables

Books (i.i.d. variables):
Borovskikh and Korolyuk (1989)
Giné and de la Peña (1999)
Dependent stationary case:
Kanagawa and Yoshihara (1994): a. s. invariance principle for completely degenerate (canonical) $U$-statistics of degree two
Aaronson, Burton, Dehling, Gilat, Hill and Weiss (1996): strong law of large numbers
Two papers by Borovkova, Burton and Dehling (2001): a version of the FCLT (along with other results)
Borisov, Volod’ko (2008): the CLT for power series’ in a weakly dependent sequence

Mixing conditions, in particular, absolute regularity are assumed; the coupling method, the method of moments e.t.c. are employed
Let for every $1 \leq p \leq \infty$ denote the projective (or maximal) tensor product

$$L_p(\Omega_1, \mathcal{F}_1, P_1) \hat{\otimes} \cdots \hat{\otimes} L_p(\Omega_d, \mathcal{F}_d, P_d).$$

Since the projective norm is stronger than the norm of $L_p(\mathcal{P}^d)$, $\hat{L}_p(\mathcal{P}^d)$ can be embedded into $L_p(\mathcal{P}^d)$.

**Example.** For $p = 2$ and $d = 2$ the space $\hat{L}_2(\mathcal{P}^2)$ can be identified with the space of (the kernels of) the trace class operators mapping $L_2(\mathcal{P})^*$ to $L_2(\mathcal{P})$.

The space $\hat{L}_p(\mathcal{P}^d)$ is preserved by the operators $(U^n, U^{*n})_{n \in \mathbb{Z}_+^d}$. We will use the denotation $(U^n, U^{*n})_{n \in \mathbb{Z}_+^d}$ for the restrictions of $(U^n, U^{*n})$ to $\hat{L}_p(\mathcal{P}^d)$ as well.
Proposition

Let \( p_1, \ldots, p_d, r \in [1, \infty] \) satisfy \( \sum_{i=1}^{d} 1/p_i = 1/r \).

Then the map sending every function

\[
(\omega_1, \ldots, \omega_d) \mapsto f_1(\omega_1) \cdots f_l(\omega_d)
\]

with \( f_1 \in L_{p_1}, \ldots, f_d \in L_{p_d} \) to the function

\[
\omega \mapsto f_1(\omega) \cdots f_d(\omega)
\]

extends in a unique way to a linear operator of norm 1

\[
D_d : L_{p_1} \hat{\otimes} \cdots \hat{\otimes} L_{p_d} \rightarrow L_r.
\]
Remark

Let \((A_n)_{n \geq 1}\) be a refining sequence of finite measurable partitions \(A_n = \{A_{1,n}, \ldots, A_{m_n,n}\}\) such that \(\mathcal{F}\) is the smallest \(\sigma\)-field containing all \(A_n, n \geq 1\). Then the operator \(D_d\) can be represented as a strong limit of the sequence of operators \((D_{d,n})_{n \geq 1}\), where

\[
D_{d,n}f = \sum_{i=1}^{m_n} \frac{I_{A_{i,n}}}{P(A_{i,n})^d} \int_{A_{i,n}^d} f(\omega_1, \ldots, \omega_d) P(d\omega_1) \cdots P(d\omega_d).
\]
Let $T_1, \ldots, T_d$ be copies of the transformation $T$ which act on $\Omega^d$ via

$$T_i(\omega_1, \ldots, \omega_i, \ldots, \omega_d) = (\omega_1, \ldots, T_i\omega_i, \ldots, \omega_d), \ i = 1, \ldots, d.$$ 

Let $\mathbb{Z}_+^d$ be the additive semigroup of $d$–tuples of nonnegative integers. The transformations $T_1, \ldots, T_d$ pairwise commute and give rise to the measure preserving action $n = (n_1, \ldots, n_d) \mapsto T^n = T_1^{n_1} \cdots T_d^{n_d}$, of $\mathbb{Z}_+^d$ on $(\Omega, \mathcal{F}, P)^d$. Set $U_k f = f \circ T_k$ for $f \in L_p$. Let $U_k^*$ be the adjoint of $U_k$ and $I$ denote the identity operator. Clearly, $U_1, \ldots, U_d$ pairwise commute, and so are $U_1^*, \ldots, U_d^*$. 

Mikhail Gordin

FUNCTIONAL LIMIT THEOREMS
From now on by a $V$-statistics of a measure preserving transformation $T$ acting on a probability space $(\Omega, \mathcal{F}, P)$ we mean the function of the form

$$\frac{1}{N^d} \sum_{1 \leq n_k \leq N, k=1,\ldots,d} D_d(h \circ T^{(n_1,\ldots,n_d)}).$$

(3)

The function $h$ is called the kernel of the corresponding $V$-statistics.
Let $T_1, \ldots, T_d$ be copies (acting on the cartesian product) of a transformation $T$. Remind that $\hat{L}_{p,\pi}(P^d) = L_{p}^\otimes d$.

**Theorem**

Let $d \geq 2$, $p \geq d$ and $r = p/d$. Let $T$ be an ergodic $P$-preserving transformation of the space $(\Omega, \mathcal{F}, P)$. Assume also that $f \in \hat{L}_{p,\pi}(P^d)$. Then, as $N \to \infty$, the sequence

$$
\frac{1}{N^d} \sum_{1 \leq n_k \leq N, k=1,\ldots,d} D_d\left( f \circ T^{(n_1,\ldots,n_d)} \right) \quad (4)
$$

converges with probability 1 and in $L_r(P)$ to the limit

$$
\int_{\Omega^d} f(\omega_1, \ldots, \omega_d) P(d\omega_1) \cdots P(d\omega_d).
$$
If $p = d$, the above Theorem applies and asserts the convergence with probability 1 and in $L_1$. 
HOEFFDING’S AND MARTINGALE-COBOUNDARY DECOMPOSITIONS
Let \((\Omega, \mathcal{F}, P)\) be a probability space and

\[
\Omega^d = \prod_{i=1}^{d} \Omega_i, \quad \mathcal{F}^d = \prod_{i=1}^{d} \mathcal{F}_i, \quad P^d = \prod_{i=1}^{d} P_i,
\]

where \(\Omega_1, \ldots, \Omega_d, \mathcal{F}_1, \ldots, \mathcal{F}_d, P_1, \ldots, P_d\) are copies of \(\Omega, \mathcal{F}\) and \(P\), respectively. Denoting by \(\pi_i\) the projection from \(\Omega^d\) onto \(\Omega_i\) \((i = 1, \ldots, d)\), we set for every \(S \in S_d\)

\[
\mathcal{F}^S = \bigvee_{i \in S} \pi_i^{-1}(\mathcal{F}_i), \quad E^S = E^{\mathcal{F}^S}, \quad \hat{E}^i = E^{\{1, \ldots, d\}\setminus\{i\}}.
\]

In other terms, \(\hat{E}^i\) integrates out the \(i\)–th variable. The identity \(I\) in \(L_p(P^d)\) decomposes as

\[
I = \prod_{i=1}^{d} (\hat{E}^i + (I - \hat{E}^i)) = \sum_{k=0}^{d} \sum_{S \in S_d^k} \prod_{i \notin S} \hat{E}^i \prod_{i \in S} (I - \hat{E}^i)
\]
For every $S \in S_d^k$ the function

$$\prod_{i \notin S} \hat{E}^i \prod_{i \in S} (I - \hat{E}^i)f$$

can be thought of as a function $f_S$ of $k$ variables $
\omega_m, m \in S$, with the property

$$\int_{\Omega} f_S(\cdots, \omega_i, \cdots) P(d\omega_i) = 0$$

for every $i \in S$. Functions of $k$ variables with this property are called **completely degenerate** or **canonical**. Observe, that for $f$ symmetric we obtain a symmetric function of $k$ variables.
The second order (compared to the SLLN) asymptotics for $V$-statistics can be studied by means of a $T$-invariant filtration and martingale approximation. We consider a (non-invertible) transformation $T$ and its canonical decreasing filtration $(T^{-n}\mathcal{F})_{n \geq 0}$. This is equivalent, up to time reversal, to considering invertible transformations, decreasing filtrations and adapted random sequences. For simplicity we assume that the transformation $T$ is exact. This means that $\bigcap_{k=0}^{\infty} T^{-k}\mathcal{F} = \mathcal{N}$, where $\mathcal{N}$ is the trivial sub $\sigma-$field of $\mathcal{F}$. 

Mikhail Gordin

FUNCTIONAL LIMIT THEOREMS
For every $k = 1, \ldots, d$, $n \geq 0$ we have

$$U_k^n U_k^* = I \text{ and } U_k^n U_k^* = E^{T_k^{-n} \mathcal{F} \times d}.$$ 

Observe that for every $1 \leq i, j \leq d$, $i \neq j$, we have

$$U_i U_j^* = U_j^* U_i.$$

Transformations $T_1, \ldots, T_d$ are completely commuting which means that they commute and enjoy the above property. The complete commutativity implies that the conditional expectations

$$(E^{T_k^{-n} \mathcal{F} \times d})_{n \geq 0, k = 1, \ldots, d}$$

commute.
For every $n = (n_1, \ldots, n_d) \in \mathbb{Z}_+^d$ we set

$$\mathcal{F}^n = T^{-n} \mathcal{F} \times d, \ E^n = E^{\mathcal{F}^n}.$$ 

Let $\overline{\mathbb{Z}_+^d} = \{0, 1, \ldots, \infty\}^d$ be a completion of $\mathbb{Z}_+^d$ endowed with the natural partial order $\leq$ which extends that of $\mathbb{Z}_+^d$. Let us extend by continuity the families $(\mathcal{F}^n)_{\mathbb{Z}_+^d}$ and $(E^n)_{\mathbb{Z}_+^d}$ to $\overline{\mathbb{Z}_+^d}$. Thus, $(\mathcal{F}^n)_{n \in \overline{\mathbb{Z}_+^d}}$ is a decreasing filtration parameterized by the partially ordered set $\overline{\mathbb{Z}_+^d}$. Let $(l, m) \mapsto l \lor m$ be the operation of taking the coordinatewise maximum in $\overline{\mathbb{Z}_+^d}$. We have $E^l E^m = E^m E^l = E^{l \lor m}$ for all $l, m \in \overline{\mathbb{Z}_+^d}$, that is the $\sigma$-fields $\mathcal{F}^l$ and $\mathcal{F}^m$ are conditionally independent given $\mathcal{F}^{l \lor m}$. 

Mikhail Gordin

FUNCTIONAL LIMIT THEOREMS
Definition

Let \((X_n, \mathcal{F}^n)_{n \in \mathbb{Z}^d_+}\) be a family of random variables defined on \((\Omega, \mathcal{F}, P)\) and sub-\(\sigma\)-fields of \(\mathcal{F}\). \((X_n, \mathcal{F}^n)_{n \in \mathbb{Z}^d_+}\) is said to be a family of reversed martingale differences if

1. the map \(\mathbb{Z}^d_+ \ni n \mapsto \mathcal{F}^n\) is decreasing (\(\mathbb{Z}^d_+\) is taken with its natural partial order, the \(\sigma\)-fields are ordered by inclusion);
2. for every \(n \in \mathbb{Z}^d_+\) the random variable \(X_n\) is measurable with respect to \(\mathcal{F}^n\);
3. \(E^{\mathcal{F}^m} X_n = 0\) whenever \(m \not\leq n\).

Variants of this definition can be found in the literature.
Let $S_d$ denote the set of all subsets of $\{1, \ldots, d\}$.

**Proposition**

Let for some $1 \leq p \leq \infty$ and $f, g \in L_p$

$$f = \left( \prod_{k=1}^{d} (I - U_k^*) \right) g.$$

Then $f$ can be represented in the form

$$f = \sum_{S \in S_d} \left( \prod_{k \in S} (U_k - I) \prod_{l \notin S} (I - U_l U_l^*) \right) h_S,$$

where for every $S \in S_d$ the function $h_S \in L_p$ is defined by

$$h_S = \left( \prod_{m \in S} U_m^* \right) g.$$
Let for some function $f$ the potential series

$$\sum_{n \in \mathbb{Z}^d_+} U^nf, \quad (7)$$

converges in the $L_p$-norm (or in $\hat{L}_p$ norm), where summation is performed over coordinate rectangles with growing edges. Then its sum presents a solution of the Poisson equation.
For a kernel $h$ of degree $d$ the following properties are equivalent:

$$E^{(n_1,\ldots,n_d)}h \xrightarrow{\max(n_1,\ldots,n_d)\to\infty} 0,$$

$$U^*(n_1,\ldots,n_d)h \xrightarrow{\max(n_1,\ldots,n_d)\to\infty} 0,$$

and

$$E^{(n_1,\ldots,n_d)}h = 0$$

whenever at least one of $n_k$ equals $\infty$. The latter property is means the canonicity of $h$.

**Canonical kernels of degree $d$ with convergent potential series form a dense subspace (among all canonical kernels of degree $d$).**
Let $\Omega = \{ z \in \mathbb{C} : |z| = 1 \}$, $P$ be the probability Haar measure on $\Omega$, $Tz = z^2$, $z \in \Omega$, $d = 2$. Clearly,

$$(Uf)(x) = f(x^2), \quad (U^*f)(x) = 1/2 \sum_{\{u:u^2=x\}} f(u).$$

If $f \in L_2(P)$ and $\int_{\Omega} f(x)P(dx) = 0$ then the series $\sum_{k \geq 0} U^*k f$ converges in $L_2$ under very mild conditions. The condition $\sum_{k \geq 0} w^{(2)}(f, 2^{-k}) < \infty$ is sufficient. Here $w^{(2)}(f, \delta)$ is the continuity modulus of $f$ in $L_2(P)$. 

Mikhail Gordin

FUNCTIONAL LIMIT THEOREMS
Let now $f_2 \in L_2(\mu^2)$ be of the form

$$f_2(x_1, x_2) = g(x_1 x_2^{-1})$$

with

$$g(x) = \sum_{k \in \mathbb{Z}} g_k x^k \in L^2(\mu).$$

Assume that $f_2 \in \hat{L}_2^{sym}$ and is canonic. This implies

$$g_0 = 0, g_{-k} = g_k, \text{ and } \sum_{k \in \mathbb{Z}} |g_k| < \infty.$$
Let $A_2$ be the Banach space of double absolutely converging Fourier series

$$a : (x_1, x_2) \mapsto \sum_{(k_1, k_2) \in \mathbb{Z}^2} a_{k_1, k_2} x_1^{k_1} x_2^{k_2}$$

furnished with the norm $| \cdot |_{A_2} : a \mapsto \sum_{(k_1, k_2) \in \mathbb{Z}^2} |a_{k_1, k_2}|$. The projective tensor norm of the space $\hat{L}_{2, \pi} \cong l_2 \hat{\otimes}_\pi l_2$ does not exceed the norm of $A_2 \cong l_1 \hat{\otimes}_\pi l_1$. Hence, the series

$$\sum_{(i_1, i_2) \in \mathbb{Z}^2_+} U^*(i_1, i_2) f_2$$

converges in $\hat{L}_{2, \pi}(\mu^2)$ if it converges in $A_2$. 
Every $U^*(i,j)$ is a contraction in $A_2$. Furthermore, 
\[ |U^*(k,0) f_2|_{A_2} = |U^*(0,k) f_2|_{A_2} = |U^k g|_{A_1}, \]
where $A_1$ is the space of one-dimensional absolutely convergent trigonometric series $a : x \mapsto \sum_{k \in \mathbb{Z}} a_k x^k$ with the norm $|a|_{A_1} = \sum_{k \in \mathbb{Z}} |a_k|$. Thus we have
\[
\sum_{(k_1,k_2) \in \mathbb{Z}^2_+} |U^*(k_1,k_2) f_2|_{A_2} \leq \sum_{0 \leq k_1 \leq k_2 < \infty} |U^*(k_1,k_2) f_2|_{A_2} + \sum_{0 \leq k_2 \leq k_1 < \infty} |U^*(k_1,k_2) f_2|_{A_2} \\
= \sum_{k \in \mathbb{Z}_+} (k + 1) (|U^*(k,0) f_2|_{A_2} + |U^*(0,k) f_2|_{A_2}) \leq 2 \sum_{k=0}^{\infty} (k + 1) |U^k g|_{A_1},
\]
(9)
Therefore, a sufficient condition for series (8) to converge in $\hat{L}_{2,\pi}(\mu^2)$ is

$$\sum_{n\in\mathbb{Z}} \sum_{k\geq 0} (k + 1)|g_{2^k n}| < \infty,$$

which holds, for example, whenever for some $C > 0$ and $\delta > 0$

$$|g_m| \leq \frac{C}{|m|(\log |m|)^{1+\delta}}, \quad m \in \mathbb{Z} \setminus \{0\}.$$
FCLT
Proposition

Let \( h \in \hat{L}_2(P^2) \) be a canonical kernel of degree 2. Assume that the limit

\[
\lim_{n_1, n_2 \to \infty} \sum_{0 \leq i_1 \leq n_1 - 1, 0 \leq i_2 \leq n_2 - 1} U^*(i_1, i_2) h
\]

exists in \( \hat{L}_2(P^2) \). Then \( h \) admits a unique representation in the form

\[
h = g + (U^{(1,0)} - I)g_1 + (U^{(0,1)} - I)g_2 + (U^{(1,0)} - I)(U^{(0,1)} - I)g_{1,2},
\]

where \( g \in \hat{L}^2(P^2), \ g_1, g_2, g_{1,2} \in \hat{L}^2(P^2) \) and

\[
E(g \mid T^{-(1,0)} \mathcal{F}^2) = 0, \ E(g \mid T^{-(0,1)} \mathcal{F}^2) = 0, \ E(g_1 \mid T^{-(1,0)} \mathcal{F}^2) = 0, \ E(g_2 \mid T^{-(0,1)} \mathcal{F}^2) = 0.
\]

Moreover, if \( h \) is a symmetric function, so is \( g \).
Assume $d = 2$. Holds for every $d \geq 1$.

**Theorem**

Let $f \in \hat{L}_2(P^2)$ be a symmetric kernel with Hoeffding's decomposition

$$f(x_1, x_2) = f_0 + f_1(x_1) + f_1(x_2) + f_2(x_1, x_2),$$

where

$$\int_X f_1(z)p(dz) = 0,$$

$f_2 \in L_p^p(\mu^2)$ and

$$\int_X f_2(z_1, x_2)\mu(dz_1) = 0.$$

Assume that the series \( \sum_{k=0}^{\infty} U^k f_1 = g_1 \) converges in \( L_2 \) and the limit

\[
\lim_{N_1,N_2 \to \infty} \sum_{(n_1,n_2)=0} U^{(n_1,n_2)} f_2
\]

exists in \( \hat{L}_2(P^2) \). Then the distributions of random variables

\[
t \mapsto N^{-3/2} \sum_{n_1,n_2=0}^{[Nt]} (f(T^{n_1}; T^{n_2} \hat{\cdot}) - f_0), \ t \in [0, 1]
\]

weakly converge to the distribution of \( 2\sigma^2_f w(\cdot) \), where \( w \) is the standard Brownian motion and

\[
\sigma^2_f = |g|^2 - |U^* g|^2.
\]
Lemma

There exists an absolute constant $C$ such that

$$
|\max_{0 \leq n_1 \leq N_1 - 1, 0 \leq n_2 \leq N_2} D_2 \sum_{(n_1, n_2) = 0} U^{(n_1, n_2)} f_2|_1 \leq C |f_2|_{2, \pi} \sqrt{N_1 N_2}
$$

(11)
H. Dehling, M. Dehling, M. Gordin. *Some limit theorems for von Mises statistics of a measure preserving transformation*. Paper in preparation.

M. Gordin. Martingale-coboundary representation for a class of random fields. *Journal of Mathematical Sciences, New York*, 163, 4, 363 – 374: 2009.