KYBERNETIKA — VOLUME *30* (1994), NUMBER 1, PAGES 77–86

THE CENTRAL LIMIT THEOREM FOR RANDOM FIELDS

MARTIN JANŽURA

The appropriate version of the central limit theorem for random fields derived from the Gibbs distributions is proved. The approach is based on a general Rosenblatt's theorem which is shown to be well suited to the case of Gibbs underlying distributions, whenever Dobrushin's uniqueness condition is satisfied.

1. INTRODUCTION

The problem of the central limit theorem (CLT) for random fields can be considered as a special case of the CLT for dependent variables. Such problem has been studied by many authors, a survey of results can be found e. g. in [1]. The crucial point of any approach consists in formulating conditions of weak dependence (such as mixing, decay of correlations, etc.).

For the random fields, i. e. the stochastic processes on a multi-dimensional integer lattice, the standard conditions cannot be, unfortunately, adapted directly, but they must be somehow modified (for the discussion cf. [6]).

For the particular case of derived random fields, i. e. those obtained by shifting a functional defined on some underlying random field, there exist long known general conditions given by [7]. Nevertheless, in spite of following the basic intention of control over the dependence structure, these conditions (cf. Theorem 2.1 below) are mostly technical and rather difficult to interpret. The aim of the present paper is to show that these conditions can be verified for the natural subclass of derived random fields with the Gibbsian type of underlying random fields.

The Gibbs random fields were originally studied in the frame of statistical mechanics. But, since they can be understood as an infinite-dimensional extension of the exponential probability distributions, their application to the problems of image processing and the statistical analysis of spatial data is straightforward (cf. e. g. [1]). Then the derived fields are consequently obtained by repeating observations of some relevant statistics.

The final result of the present paper, i. e. the appropriate version of the CLT, obeys a natural interpretation of its assumptions. Moreover, it can be considered as a generalization of Theorem 4.1 in [5], which is obtained on the basis of a different version of the general CLT, and which contains two moment type conditions. In the present version (cf. Theorem 4.8 below) one of the conditions is relaxed and the other is completely removed. The main results of the paper are given in Section 4, while Sections 2 and 3 contain definitions and some preliminary results.

2. CENTRAL LIMIT THEOREM FOR RANDOM FIELDS

Let $T = Z^d$ denote the d-dimensional lattice. For a given underlying random field (r. f.)

$$(\Omega, \mathcal{A}, P, Y = \{Y_t : (\Omega, \mathcal{A}) \longrightarrow (X, F)\}_{t \in T})$$

and a real-valued Borel-measurable functional

$$f : (X^T, F^T) \longrightarrow (R, \mathcal{B}(R))$$

we define the derived r. f. by

$$\left(X^T, F^T, \mu = PY^{-1}, Y^f = \left\{ Y_t^f = f \circ \theta_t : (X^T, F^T) \to (R, \mathcal{B}(R)) \right\}_{t \in T} \right),$$

where θ_t is the corresponding shift for every $t \in T$, i. e. the transform on X^T defined through

$$[\theta_t(x)]_s = x_{t+s} \qquad \text{for every } s \in T,\ x \in X^T.$$

The r. f. Y is called stationary if its distribution μ is shift invariant, i. e.

$$\mu\, \theta_t^{-1} = \mu \qquad \text{for every } t \in T.$$

If moreover $\mu(B) = 1$ for every $B \in F^T$ satisfying $\mu(B) > 0$ and $B = \theta_t^{-1} B$ for every $t \in T$, then the r. f. Y is called ergodic.

For every $S \subset T$ we denote $\mathcal{F}^S = \mathrm{Pr}_S^{-1} F^S \subset F^T$, where $\mathrm{Pr}_S : X^T \to X^S$ is the corresponding projection function.

Further, let us denote by $L^2(S) = L^2(X^T, \mathcal{F}^S, \mu)$ the space of square integrable $\mathcal{F}^S$-measurable functions.

Let us provide the lattice T with the lexicographical ordering "$\prec$". For every $t \in T$ let t^- denote its immediate predecessor. We set

$$T(t^-) = \{s \in T;\ s \prec t\}, \qquad T(t) = T(t^-) \cup \{t\}$$

and define

$$H_t = L^2\left(T(t)\right) \ominus L^2(T(t^-)).$$

Thus, H_t is the subspace of $L^2(T)$ given by the orthogonal complement of $L^2(T(t^-))$ in $L^2(T(t))$.

For every $f \in L^2(T)$ and positive integer k let us denote

$$f^{(k)} = \mathrm{Proj}\left(\bigoplus_{t \in V^k} H_t \right)(f)$$

where Proj(H) means the projection onto the subspace H, and

$$V^k = \{t \in T;\ |t_i| \le k \text{ for every } i = 1, \dots, d\}$$

is the corresponding cube.

Similarly, we denote $V(a) = \{t \in T;\ 0 \le t_i < a_i \text{ for every } i = 1, \dots, d\}$ for every $a \in T$ and we write $a \to \infty$ for $\min\{a_i;\ i = 1, \dots, d\} \to \infty$.

We say that the central limit theorem (CLT) holds for the derived r.f. Y^f if there exists $\sigma^2 \in [0, \infty)$ such that

$$|V(a)|^{-\frac{1}{2}} \sum_{t \in V(a)} \left(Y_t^f - E_\mu Y_t^f\right) \Longrightarrow N(0, \sigma^2) \qquad \text{in distribution } [\mu] \text{ for } a \to \infty,$$

where $N(0, \sigma^2)$ stands for the normal distribution with the zero mean value and the variance equal to σ^2. (For $\sigma^2 = 0$ we deal with the distribution concentrated to the zero point.)

Theorem 2.1. (Rosenblatt [7]) Under the assumptions

i) Y is ergodic,

ii) $f \in L^2(T)$,

iii) $\sum_{t \in T} \left|\operatorname{cov}\left(Y_t^f, Y_0^f\right)\right| < \infty$,

iv) $Y_0^f - EY_0^f \in \bigoplus_{t \in T} H_t$,

v) $\sum_{t \in T} \left|\operatorname{cov}\left(Y_0^f - (Y_0^f)^k, \left[Y_0^f - (Y_0^f)^k\right] \circ \theta_t\right)\right| \longrightarrow 0$ for $k \to \infty$,

the CLT holds for the derived r.f. Y^f with

$$\sigma^2(Y^f) = \sum_{t \in T} \operatorname{cov}\left(Y_t^f, Y_0^f\right).$$

The proof for $d \le 2$ is given in [7] (Theorem 3.2) but the generalization for higher dimensions is straightforward.

3. GIBBS RANDOM FIELDS

Let us denote by $\mathcal{A}$ the system of all subsets of T, and by $\mathcal{K}$ the system of finite non-void subsets.

A specification is a class of probability kernels

$$\Pi = \left\{\Pi_V(\cdot|\cdot) : \ F^T \times X^T \to [0, 1]\right\}_{V \in \mathcal{K}}$$

satisfying the following conditions:

a) $\Pi_V(B|\cdot)$ is $\mathcal{F}^{T \setminus V}$-measurable for every $B \in \mathcal{F}^T$,

b) $\Pi_V(B|\cdot) = \chi_B$ for every $B \in \mathcal{F}^{T\setminus V}$,

c) $\Pi_V(B|\cdot) = \int \Pi_W(B|x)\,\Pi_V(dx|\cdot)$ for every $B \in \mathcal{F}^T$ and $V,\, W \in \mathcal{K},\; W \subset V$.

A specification is stationary if

$$\Pi_V(B|x) = \Pi_{V-t}(\theta_t B|\theta_t x)$$

for every $V \in \mathcal{K},\; t \in T,\; B \in \mathcal{F}^T,\; x \in X^T$.

A probability measure μ defined on (X^T, F^T) is called the Gibbs distribution with respect to the specification Π (we write $\mu \in G(\Pi)$) if

$$E_\mu\left[\chi_B \,|\, \mathcal{F}^{T\setminus V}\right](\cdot) = \Pi_V(B|\cdot)$$

holds a.s. $[\mu]$ for every $V \in \mathcal{K}$ and $B \in F^T$. A r.f. is called Gibbs r.f. if its distribution is a Gibbs distribution.

For $a,\, b \in T$ we define

$$\gamma_{ab} = \frac{1}{2}\sup\left\{\left\|\Pi_{\{a\}}^{\{a\}}(\cdot|x) - \Pi_{\{a\}}^{\{a\}}(\cdot|z)\right\|_{TV};\quad x,\, z \in X^T,\; x_t = z_t \text{ for } t \neq b\right\}$$

where the total variation is meant by the norm, and Π_V^W for $W \subset T$ is the restriction of Π_V to the σ-algebra $\mathcal{F}^W$.

If $\sum_{b\in T}\gamma_{ab} \leq \gamma < 1$ for every $a \in T$, the specification is said to satisfy Dobrushin's condition.

For a stationary specification we may write $\gamma_{ab} = \gamma_{a-b}$, and the Dobrushin's condition may be rewritten in the form

$$\sum_{a\in T}\gamma_a = \gamma < 1.$$

From now let us suppose (X, F) to be a compact metric space with its Borel σ-algebra, and denote by $C(X^T)$ the space of continuous functions on the product compact space X^T. Moreover, for $f \in C(X^T)$ and $s \in T$ we set

$$\varphi_s(f) = \sup\left\{|f(x) - f(z)|,\; x_t = z_t \text{ for } t \neq s\right\},$$

and denote

$$C^\varphi = \left\{f \in C(X^T);\; \varphi(f) = \sum_{s\in T}\varphi_s(f) < \infty\right\}.$$

Note that C^φ is a dense subset of $C(X^T)$.

A specification Π is said to be continuous if

$$\int f(x)\,\Pi_V(dx|\cdot) \in C(X^T)$$

holds for every $f \in C(X^T)$ and $V \in \mathcal{K}$.

Let us emphasize that if a continuous specification Π satisfies Dobrushin's condition then the Gibbs distribution μ is uniquely determined, i.e. $G(\Pi) = \{\mu\}$ (cf. e.g. Corollary 2.3 in [5]). If the specification Π is stationary then the unique Gibbs distribution μ is stationary and ergodic.

Proposition 3.1. A continuous specification satisfying Dobrushin's condition can be uniquely extended to a system of kernels

$$\overline{\Pi} = \left\{ \Pi_S(\cdot|\cdot): \; F^T \times X^T \to [0,1] \right\}_{S \in \mathcal{A}}$$

with the same properties, i.e. a), b), c) in the definition of specification remain true with $\mathcal{A}$ instead of $\mathcal{K}$, and $\int f(x)\,\Pi_S(dx|\cdot) \in C(X^T)$ equals to $E_\mu\left[f \,|\, \mathcal{F}^{T\setminus S}\right](\cdot)$ a.s. $[\mu]$ for every $f \in C(X^T)$.

Proof. Following Section 3 in [4] we observe that for every $\emptyset \neq S \subset T$ the conditional distribution

$$\mu\left(\cdot|\mathcal{F}^{T\setminus S}\right)(x)$$

exists, being a.s. $[\mu]$ equal to $\Pi_S(\cdot|x)$ which is the unique Gibbs distribution with respect to the specification

$$\Pi^{S,x} = \left\{ \Pi_V^{S,x}(\cdot|\cdot, x_{T\setminus S}) \right\}_{V \in \mathcal{K},\, V \subset S}$$

(for $\Pi^{S,x}$ the Dobrushin's condition is satisfied as well).

Due to Corollary 2.4 in [5]

$$\int f(x)\,\Pi_S(dx|\cdot)$$

is a uniform limit of $\int f(x)\,\Pi_V(dx|\cdot)$ for $V \nearrow S$ and $f \in C^\varphi$.

Since C^φ is dense in $C(X^T)$ it holds

$$\int f(x)\,\Pi_S(dx|\cdot) \in C(X^T)$$

for every $f \in C(X^T)$.

For $S = \emptyset$ we simply set $\Pi_\emptyset$ equal to the indicator function. □

4. CLT FOR GIBBS RANDOM FIELDS

Let us suppose that (X, F) is a compact metric space with its Borel σ-algebra and the distribution μ of the underlying r.f. Y is the Gibbs distribution with respect to a continuous and stationary specification Π that satisfies Dobrushin's condition. Note that the condition i) of Theorem 2.1 is immediately satisfied. Before formulating the main theorem we introduce several auxiliary results.

For the sake of brevity we shall write $f_{T\setminus S}$ instead of

$$\int f(x)\,\Pi_S(dx|\cdot)$$

for every $f \in C(X^T)$ and $S \in \mathcal{A}$.

For every $S \in \mathcal{A}$ and $f \in C^{\varphi}$ we denote

$$K_f(S) = \sum_{c \in S} \sum_{b \in T} \chi_{cb} \varphi_b(f)$$

where

$$\chi = \sum_{n=0}^{\infty} \Gamma^n, \qquad \Gamma = \{\gamma_{ab}\}_{a,b \in T}.$$

Note that

$$K_f(S) \leq K_f(T) = (1-\gamma)^{-1} \varphi(f) < \infty,$$

and

$$K_{f \circ \theta_t}(S) = K_f(S - t)$$

for every $S \in \mathcal{A}$ and $t \in T$.

Lemma 4.1. Let $f \in C^{\varphi}$.

i) Then $f_S \in C^{\varphi}$ for every $S \in \mathcal{A}$. Namely

$$\varphi_t(f_S) \leq K_f(\{t\}) \quad \text{for every } t \in T \quad \text{and therefore } \varphi(f_S) \leq K_f(T).$$

ii) For every $S_1 \subset S_2$; $S_1, S_2 \in \mathcal{A}$ it holds

$$\|f_{S_1} - f_{S_2}\| \leq K_f(S_2 \setminus S_1).$$

iii) Let also $g \in C^{\varphi}$. Then

$$|\operatorname{cov}(f, g)| \leq \sum_{s \in T} K_f(\{s\}) \, K_g(\{s\}).$$

Proof. The statements i) and ii) follow from Corollary 2.4 in [5] while the statement iii) follows from Theorem 3.2 [5]. □

Corollary 4.2. Let $f \in C^J$. Then $\varphi(f - f_{V^k}) \to 0$ for $k \to \infty$.

Proof. Let us fix k^0. Then it holds

$$\begin{aligned} \varphi(f - f_{V^k}) &\leq 2|V^{k^0}| \, \|f - f_{V^k}\| + \sum_{s \in T \setminus V^{k^0}} \varphi_s(f) + \varphi_s(f_{V^k}) \leq \\ &\leq 2|V^{k^0}| \, K_f(T \setminus V^k) + 2K_f(T \setminus V^{k^0}) \end{aligned}$$

by i) and ii) of Lemma 4.1.

Thus, for $k \to \infty$ and subsequently for $k^0 \to \infty$ we obtain the result since obviously $K_f(T \setminus V^k) \to 0$ for $k \to \infty$. □

Proposition 4.3. Suppose $f,\ f_n \in C^\varphi$ and $\varphi(f_n - f) \to 0$ for $n \to \infty$. If the CLT holds for every derived r.f. Y^{f_n} then it holds also for Y^f.

P r o o f. Due to the well-known theorem (cf. e. g. [2], Theorem 4.2) it is sufficient to prove that

$$\left|\sigma^2(Y^{f_n}) - \sigma^2(Y^f)\right| \qquad \text{for } n \to \infty$$

and

$$\sum_{t\in T}\left|\operatorname{cov}\left(Y_t^{f_n - f},\ Y_0^{f-f}\right)\right| \longrightarrow 0 \qquad \text{for } n \to \infty.$$

But, according to Lemma 4.1 iii), the first expression may be bounded e. g. by

$$(1-\gamma)^{-2}\ [2\varphi(f) + \varphi(f - f_n)]\ \varphi(f - f_n)$$

and the second one by

$$(1-\gamma)^{-2}\ [\varphi(f - f_n)]^2\,.$$

□

Remark 4.4. From Proposition 4.3 and Corollary 4.2 it follows that if we prove the CLT for all "local (cylinder) functions" from C^φ, i.e. those measurable with respect to some $\mathcal{F}^S$, $S \in \mathcal{K}$, then the CLT remains valid for every $f \in C^\varphi$.

Since it is well known that

$$\operatorname{Proj}(H_t) f = E_\mu\left[f|\mathcal{F}^{T(t)}\right] - E_\mu\left[f|\mathcal{F}^{T(t^-)}\right] \qquad \text{a.s. } [\mu]$$

we may fix

$$\operatorname{Proj}(H_t) f = f_{T(t)} - f_{T(t^-)} \qquad \text{for every } t \in T \text{ and } f \in C(X^T),$$

and, consequently,

$$f^{(k)} = \sum_{t\in V^k}\left[f_{T(t)} - f_{T(t^-)}\right] = \sum_{r\in V^k_{d-1}}\left[f_{T(r,k)} - f_{T(r,-k-1)}\right].$$

By re-arranging we also obtain

$$f^{(k)} = f_{T(k_d)} - \sum_{r\in V^k_{d-1}\setminus\{-k_{d-1}\}}\left[f_{T(r,-k-1)} - f_{T(r^-,k)}\right] - f_{T(-k_{d-1},-k-1)}$$

where $k_d = \{k,\dots,k\} \in Z^d$.

Proposition 4.5. Let $f \in C^\varphi$. If

$$\varphi\left(f - E_\mu f - f^{(k)}\right) \longrightarrow 0 \qquad \text{for } k \to \infty$$

then the CLT holds for the derived r.f. Y^f.

P r o o f. For $f^{(k)}$ the assumptions iv) and v) of Theorem 2.1 are easily satisfied, while the assumptions ii) and iii) hold for every $f \in C^\varphi$ (thanks to Lemma 4.1 iii)).

Since $f^{(k)} \in C^\varphi$ for every k by Lemma 4.1 i), the statement follows from Proposition 4.3. □

Proposition 4.6. Let $|V^k|\cdot K_f(T\backslash V^k) \to 0$ for $k \to \infty$, then $\varphi\left(f - E_\mu f - f^{(k)}\right) \to 0$ for $k \to \infty$.

Proof. We have

$$f - E_\mu f - f^{(k)} = f - f_{T(k_d)} + \\ + \sum_{r \in V_{d-1}^k \setminus \{-k_{d-1}\}} \left[f_{T(r,-k-1)} - f_{T(r-,k)}\right] + f_{T(-k_{d-1},-k-1)} - f_\emptyset,$$

and therefore

$$\left\|f - E_\mu f - f^{(k)}\right\| \le K_f(T \setminus V^k) \qquad \text{by Lemma 4.1 ii)}$$

and

$$\varphi_s\left(f - E_\mu f - f^{(k)}\right) \le \left(2|V_{d-1}^k| + 1\right)\, K_f(\{s\}) \qquad \text{by Lemma 4.1 i).}$$

Thus, using the estimate

$$\varphi\left(f - E_\mu f - f^k\right) \le 2 \cdot |V^k| \cdot \|f - E_\mu f - f^k\| + \sum_{s \notin V^k} \varphi_s(f - E_\mu f - f^{(k)})$$

we conclude

$$\varphi(f - E_\mu f - f^k) \le 3|V^k|\, K_f(T \setminus V^k)$$

which yields the statement. □

By $\|t\|$ for $t \in T$ we denote the norm given by $\|t\| = \max_{i=1,\dots d} |t_i|$.

Proposition 4.7. Let $f \in C^\varphi$ be $\mathcal{F}^W$-measurable, $W \in \mathcal{K}$, and

$$\sum_{t \in T} \|t\|^d \gamma_t = C_\gamma < \infty.$$

Then

$$|V^k|\, K_f(T \setminus V^k) \longrightarrow 0 \qquad \text{for } k \to \infty.$$

Proof. Since $|V^k|(2k+1)^d \le 3^d k^d$ and $k \le \|t\|$ for $t \in T \setminus V^k$ we may write

$$|V^k|\, K_f(T \setminus V^k) \le 3^d \sum_{t \in T \setminus V^k} \|t\|^d \sum_{b \in W} \chi_{tb} \varphi_b(f).$$

Therefore, it is sufficient to prove

$$B = \sum_{t \in T} \|t\|^d \chi_{t0} < \infty.$$

Let us denote $\chi^{(n)} = \sum_{i=0}^n \Gamma^i$ and $B^n = \sum_{t \in T} \|t\|^d \chi_{t0}^{(n)}$. We obtain

$$B^n = \sum_{t,} \sum_{s \in T} \|t + s\|^d \chi_{t0}^{(n-1)} \gamma_s \le \left[\left(\gamma\, B^{n-1}\right)^{\frac{1}{d}} + \left(\frac{1-\gamma^n}{1-\gamma}\, C_\gamma\right)^{\frac{1}{d}}\right]^d$$

by the triangular inequality of the ℓ_d-norm. Thus $B = \lim_{n\to\infty} B_n < \infty$. □

Now we may formulate the appropriate version of the CLT which is the main result.

Theorem 4.8. Let the underlying Gibbs distribution satisfy moreover

$$\sum_{t\in T} \|t\|^d \gamma_t < \infty.$$

Then the CLT holds for every Y^f, $f \in C^\varphi$.

Proof. The statement follows directly from the preceding propositions. □

Remark 4.9. For Markov random fields the additional assumption is easily satisfied since then $\gamma_t = 0$ for $t \in T \setminus V$ with some $V \in \mathcal{K}$.

5. CONCLUDING REMARK

The assumption of compact metric space X was made not to complicate the proofs too much. But, the substantial estimates contained in Lemma 4.1 can be obtained, with the aid of the Vasershtein distance, for non-compact case as well. Therefore even the main result can be with some technical difficulties generalized.

For the image processing or the statistical analysis of spatial data (cf. [1]) we usually consider the specifications to be given by a potential (cf. e.g. [5]). The potential is a collection of functions which offers our equivalent description of a Gibbs random field but it can be directly understood as a parametrization. Thus the problem of parameter estimation for Gibbs random fields can be formulated. For investigating the asymptotic properties, namely the asymptotic normality, of some relevant estimators, the validity of the CLT for appropriate derived random fields is crucial. Since the Markov property of the underlying random field can be assumed without any major restriction, the additional condition in Theorem 4.8 does not cause any practical problem. On the other hand the Dobrushin's uniqueness condition is really substantial and cannot be easily avoided.

(Received October 20, 1992.)

REFERENCES

[1] J. Besag: On the statistical analysis of dirty pictures. J. Roy. Statist. Soc. B *48* (1986), 3, 259–302.

[2] P. Billingsley: Convergence of Probability Measures. Wiley, New York 1968.

[3] R. L. Dobrushin and B. Tirozzi: The central limit theorem and the problem of equivalence of ensembles. Commun. Math. Phys. *54* (1977), 174–192.

[4] H. Föllmer: On the global Markov property. In: Quantum Fields-Algebras Processes (L. Streit, ed.) Springer, Wien 1980.

[5] H. Künsch: Decay of correlations under Dobrushin's uniqueness condition and its applications. Commun. Math. Phys. *84* (1982), 207–222.

[6] B. S. Nahapetian: Limit theorems and statistical mechanics. In: Random Fields (J. Fritz, J. L. Lebowitz, D. Szasz, eds.), Colloquia mathematica societatis Janos Bolyai, *27* (1981), 799–820.
[7] H. Rosenblatt: Central limit theorem for stationary processes. Proc. 6th Berkeley Symposium, Vol. II (1970), 551–561.

RNDr. Martin Janžura, CSc., Ústav teorie informace a automatizace AV ČR (Institute of Information Theory and Automation – Academy of Sciences of Czech Republic), Pod vodárenskou věží 4, 182 08 Praha. Czech Republic.