HAL
open science

Asymptotic properties of the residuals process for stationary marked Gibbs point processes

Jean-François Coeurjolly

To cite this version:

Jean-François Coeurjolly. Asymptotic properties of the residuals process for stationary marked Gibbs point processes. 41èmes Journées de Statistique, Société Française de Statistique, May 2009, Bordeaux, France. inria-00386660

HAL Id: inria-00386660

https://inria.hal.science/inria-00386660

Submitted on 22 May 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Asymptotic properties of the residuals process for stationary marked Gibbs point processes

J.-F. Coeurjolly

Sagag Team, Laboratory Jean Kuntzmann/ CNRS,
BSHM, 1251 Av. Centrale, Grenoble Cedex 09 (France)

Résumé

Dans le contexte des processus ponctuels spatiaux, la notion de résidus a été proposée très récemment par Baddeley et al. (2005). Elle fournit un outil de diagnostic très riche pour juger de la qualité d'ajustement d'un modèle paramétrique de processus ponctuel. Ces résidus sont une extension de ceux définis en dimension un pour des processus de comptage et sont basés sur des versions empiriques des deux termes apparaissant dans la célèbre Formule de Georgii-Nguyen-Zeissin. Cet exposé présentera les propriétés asymptotiques du processus des résidus pour des processus ponctuels marqués de Gibbs stationaires. En particulier, nous obtenons la consistance et la normalité asymptotique pour une large classe de résidus, incluant notamment ceux définis dans Baddeley et al. (2005) et partiellement étudiés dans Baddeley et al. (2008) (résidus bruts, résidus inversés, résidus de Pearson). Nous généralisons ce résultat pour proposer un test d'adéquation basé sur des estimations fonctionnelles de la fonction d'espace vide qui caractérise mieux la distribution d'un processus ponctuel spatial qu'un seul type de résidus.

Abstract

In the context of spatial point processes, the notion of residuals has been proposed very recently by Baddeley et al. (2005). It provides a very rich diagnostic tool to investigate the quality of adjustment of a parametric spatial point process. Residuals are an extension of the ones available in one dimension for counting processes and are based on empirical versions of both terms appearing in the well-known Georgii-Nguyen-Zeissin Formula. This talk focuses on asymptotic properties of the residuals process for stationary marked Gibbs point processes. In particular, the consistency and the asymptotic normality are obtained for a wide class of residuals including the classical ones proposed by Baddeley et al. (2005) and partially studied by Baddeley et al. (2008) (raw residuals, inverse residuals, Pearson residuals). We also generalize this result for defining a goodness-of-fit test for these processes based on functionals estimations of the empty space function which better characterizes the distribution of a spatial point process rather than one type of residuals.

1 Background, examples and definitions

Let us start with some general notation on point processes. The data consist in a countable subset $\varphi = \{x_1^{m_1}, \ldots, x_n^{m_n}\}$ (where n is not fixed), called a configuration of marked points ($x_i \in \mathbb{R}^2$ (for the sake of simplicity) are location and $m_i \in \mathbb{M}$ are marks). The state space will de denoted by $\mathbb{S} = \mathbb{R}^2 \times \mathbb{M}$. Let λ^2 denote the Lebesgue measure on $\mathbb{R}^2$, λ^{m} the probability mark measure and let $\mu := \lambda^2 \otimes \lambda^{\mathrm{m}}$. Finally denote by $\varphi_\Lambda := \varphi \cap (\Lambda \cap \mathbb{M})$, $\Lambda \subset \mathbb{R}^2$ realization of the variable Φ_Λ.

The (marked) Poisson point process constitutes the reference model for spatial point processes and is related to the notion of independence. The realization of a Gibbs point process is the realization of some variable with conditional density (conditionnally on some fixed outside configuration φ^o) with respect to the Poisson process expressed as

$$f_{\Phi_\Lambda}^{\Phi_{\Lambda^c}=\varphi_{\Lambda^c}^o}(\varphi_\Lambda) \propto \exp\left(-V\left(\varphi_\Lambda | \varphi_{\Lambda^c}^o\right)\right)$$

where $V(\varphi_\Lambda|\varphi^o_{\Lambda^c}) := V(\varphi_\Lambda \cup \varphi^o_{\Lambda^c}) - V(\varphi^o_{\Lambda^c})$ measures the energy to insert the configuration φ_Λ into $\varphi^o_{\Lambda^c}$. In the rest of the paper, the framework is restricted to stationary marked Gibbs point processes based on energy function $V(\varphi;\theta)$, invariant by translation and parametrized in an exponential form by some $\theta \in \Theta$ where Θ is some compact set of $\mathbb{R}^p$ $(p \geq 1)$, i.e.

$$V(\varphi;\theta) := \theta^T \mathbf{v}(\varphi), \quad \text{where } \mathbf{v}(\varphi) = (v_1(\varphi), \ldots, v_p(\varphi)).$$

Models may be defined by defining through the local energy to insert a marked point x^m into the configuration φ

$$V(x^m|\varphi;\theta) := V(\varphi \cup \{x^m\};\theta) - V(\varphi;\theta) = \theta^T \mathbf{v}(x^m|\varphi).$$

Let us cite some well-known examples :

- Multi-type Poisson point process : $\mathbb{M} = \{1, \ldots, M\}$

$$V(\varphi;\theta) = \sum_{m=1}^{M} \theta^m |\varphi^m|, \qquad V(x^m|\varphi;\theta) = \theta^m.$$

- Multi-type Strauss point process : $\mathbb{M} = \{1, \ldots, M\}$

$$V(\varphi;\theta) = \sum_{m=1}^{M} \theta_1^m |\varphi^m| + \sum_{m=1}^{M} \theta_2^m \sum_{\{x^m, y^m\} \in \mathcal{P}_2(\varphi)} \mathbf{1}_{[D_1^m, D_2^m]}(||y - x||).$$

  Alternatively

$$V(x^m|\varphi;\theta) = \sum_{m=1}^{M} \theta_1^m + \sum_{m=1}^{M} \theta_2^m \sum_{y^m \in \varphi} \mathbf{1}_{[D_1^m, D_2^m]}(||y - x||).$$

  This process is defined for $\theta_2^m \geq 0$ and $D_1^m = 0$ (inhibition assumption) or when $\theta_2^m \in \mathbb{R}$ and $D_1^m = \delta > 0$ (hard-core assumption).

- Area interaction point process : $\mathbb{M} = \{0\}$, $R > 0$

$$V(\varphi;\theta) = \theta_1 |\varphi| + \theta_2 \left| \cup_{x \in \varphi} \mathcal{B}(x, R) \right|,$$

  where $\mathcal{B}(x, R)$ is the ball centered at x with radius R.

Concerning such models, the first statistical challenge is to estimate the true value of the parameter vector ($\theta^\star$ assumed to be in $\mathring{\Theta}$). This may be done via the maximisation of the likelihood (MLE) or the pseudo-likelihood (MPLE). The last one is an excellent alternative to the MLE since it avoids the computation of the normalizing constant which

is particularly costly (see the next section for an accurate definition and Billiot, Coeurjolly and Drouilhet (2008) for asymptotic results). Once you have estimated the parameter vector of interest, the second challenge is to quantify the quality of adjusment of the model to data. Baddeley et *al.* (2005) defined a large class of such measures based on the well-known Georgii-Nguyen-Zeissin Formula, expressed here for stationary marked Gibbs point processes

Theorem 1 (Georgii-Nguyen-Zeissin Formula) *For any function $h(\cdot,\theta) : \Omega_f \to \mathbb{R}$ such that the following quantities are finite*

$$E\left(h\left(0^M|\Phi;\theta^\star\right)e^{-V\left(0^M|\Phi;\theta^\star\right)}\right) = E\left(h\left(0^M f|\Phi\setminus 0^M;\theta^\star\right)\right) \tag{1}$$

where for any $\theta \in \Theta$, $(m,\varphi) \in \mathbb{M}\times\Omega_f$, $h(x^m|\varphi;\theta) := h(\varphi\cup x^m;\theta) - h(\varphi;\theta)$ and where $M \sim \lambda^{\mathrm{m}}$

Now define the innovations and the residuals process based on empirical versions of both terms appearing in (1).

Definition 1 *For any bounded domain Λ, let us define*

- *Innovations process:*

$$I_\Lambda(\varphi;h,\theta^\star) := \frac{1}{|\Lambda|}\int_{\Lambda\times\mathbb{M}} h(x^m|\varphi;\theta^\star)e^{-V(x^m|\varphi;\theta^\star)}\mu(dx^m) - \frac{1}{|\Lambda|}\sum_{x^m\in\varphi_\Lambda} h(x^m|\varphi\setminus x^m;\theta^\star),$$

- *Residuals process: let $\widehat{\theta}_n(\varphi)$ be an estimate of $\theta^\star$ based on φ_Λ*

$$R_\Lambda(\varphi;h) := \frac{1}{|\Lambda|}\int_{\Lambda\times\mathbb{M}} h(x^m|\varphi;\widehat{\theta}_n(\varphi))e^{-V\left(x^m|\varphi;\widehat{\theta}_n(\varphi)\right)}\mu(dx^m) - \frac{1}{|\Lambda|}\sum_{x^m\in\varphi_\Lambda} h(x^m|\varphi\setminus x^m;\widehat{\theta}_n(\varphi))$$

Clearly, the last notion is the most interesting as a practical point of view. Let us describe the main examples considered by Baddeley et *al.* (2005) (in the context of stationary point processes).

1. $h(x^m|\varphi,\theta) = 1$: raw innovations and residuals
2. $h(x^m|\varphi,\theta) = e^{V(x^m|\varphi;\theta)}$: inverse innovations and residuals.
3. $h(x^m|\varphi,\theta) = e^{V(x^m|\varphi;\theta)/2}$: Pearson innovations and residuals.

In particular, one may note that the raw residuals is a difference of two estimates of the intensity of the point process : the first one is a parametric one and depends on the model while the second one is a non parametric one (since it equals to $|\varphi_\Lambda|/|\Lambda|$).

2 Asymptotic results of the residuals process

In order to present asymptotic results, let us first consider general assumptions, denoted by [Mod], on our models:

1. [Mod:S] *Stability of the local energy*: $\exists K \geq 0$, $\forall (m, \varphi) \in \mathbb{M} \times \Omega_f$
$$V\left(0^m|\varphi;\theta\right) \geq -K.$$
2. [Mod:L] *Locality of the local energy*: $\exists D \geq 0$, $\forall (m, \varphi) \in \mathbb{M} \times \Omega_f$
$$V\left(0^m|\varphi;\theta\right) = V\left(0^m|\varphi_{B(0,D)};\theta\right).$$
3. [Mod:I] *Integrability*: $\exists \kappa_i^{(\sup)} \geq 0$, $k_i \in \mathbb{N}$, $\forall (m, \varphi) \in \mathbb{M} \times \Omega_f$
$$v_i(0^m|\varphi) \leq \kappa_i^{(\sup)} |\varphi_{\mathcal{B}(0,D)}|^{k_i}, \quad i = 1, \ldots, p.$$
4. [Mod:Id] *Identifiability condition*:: for all $\theta \in \Theta \setminus \theta^\star$,
$$P\left(V\left(0^M|\Phi;\theta\right) \neq V\left(0^M|\Phi;\theta^\star\right)\right) > 0$$

The assumptions [Mod:S] and [Mod:L] are quite well-known in the context of Gibbs point processes. In particular, they ensure the existence of stationary measures. These assumptions include a very large class of known parametric models of marked point processes : Poisson process, multi-type Poisson process, multi-strauss marked point process, Strauss-disc type process, Geyer's triplet point process, Area interaction point process,.... It also include many models where the interaction is no more measured on the complete graph ($\mathcal{P}_2(\varphi)$) but on some more structured one such as the k-nearest neighbour graph or the Delaunay graph (*e.g.* [3]). Some of these models are known under the names of Ord's model where the energy depends the area of Voronoï cells.

In the rest of the paper, the data is assumed to be observed in the domain $\Lambda_n \oplus D^\vee$ (where $D^\vee \geq D$), where the domain $\Lambda_n = [-n, n]^2$ (for the sake of simplicity). Moreover, the estimate appearing in the definition of the residuals will be the maximum pseudo-likelihood estimate, i.e.
$$\widehat{\theta}_n(\varphi) = \widehat{\theta}_n^{MPLE}(\varphi) := \mathrm{argmax}_{\theta\in\Theta} LPL_{\Lambda_n}\left(\varphi;\theta\right)$$
where the log-pseudo likelihood on the domain Λ_n, $LPL_{\Lambda_n}\left(\varphi;\theta\right)$ is defined by
$$LPL_{\Lambda}\left(\varphi;\theta\right) = -\int_{\Lambda\times\mathbb{M}} e^{-V(x^m|\varphi;\theta)}\mu(dx^m) - \sum_{x^m\in\varphi_\Lambda} V\left(x^m|\varphi\setminus x^m;\theta\right),$$
see Billiot, Coeurjolly and Drouilhet (2008) for more details and asymptotic results (consistency and asymptotic normality). Finally, consider the following assumptions denoted by [H] and related to the analysing function h:

1. $h(\cdot;\theta)$ looks like $V(\cdot;\theta)$ (invariance by translation, locality).
2. Assume that for all $(m,\varphi)\in\mathbb{M}\times\Omega_f$, $h(0^m|\varphi;\cdot)$ is differentiable for $\theta\in\mathcal{V}(\theta^\star)$ and assume that for $j=1,\dots,p$ and for all $\theta\in\mathcal{V}(\theta^\star)$

$$E\left(|h(0^M|\Phi;\theta)|^3 e^{-V\left(0^M|\Phi;\theta\right)}\right)<+\infty \quad \text{and} \quad E\left(\left|\frac{\partial h}{\partial\theta_j}(0^M|\Phi;\theta)\right| e^{-V\left(0^M|\Phi;\theta\right)}\right)<+\infty.$$

Now we are ready to present our main result concerning the innovations and the residuals processes

Theorem 2 *Under the assumptions [Mod] and [H] and under the true model,*

(i) for $P_{\theta^\star}-$a.e. φ, $I_{\Lambda_n}(\varphi;h,\theta^\star)$ and $R_{\Lambda_n}(\varphi;h)$ converge towards 0, as $n\to+\infty$.

(ii) the following convergences hold in distribution, as $n\to+\infty$

$$\begin{aligned}
|\Lambda_n|^{1/2}\, I_{\Lambda_n}(\Phi;h,\theta^\star) &\to \mathcal{N}\left(0,\sum_{i\in\mathbb{B}(0,\lceil D\rceil)} E\left(I_{\Delta_0}(\Phi;h,\theta^\star)I_{\Delta_i}(\Phi;h,\theta^\star)\right)\right)\\
|\Lambda_n|^{1/2}\, R_{\Lambda_n}(\Phi;h) &\to \mathcal{N}\left(0,\sum_{i\in\mathbb{B}(0,\lceil D\rceil)} E\left(W_{\Delta_0}(\Phi;h,\theta^\star)W_{\Delta_i}(\Phi;h,\theta^\star)\right)\right)
\end{aligned}$$

where

$$W_{\Delta_i}(\varphi;h,\theta^\star):=I_{\Delta_i}(\varphi;h,\theta^\star)+\underline{\mathbf{U}}^{(2)}(\theta^\star)^{-1}\mathbf{LPL}^{(1)}_{\Delta_i}(\varphi;\theta^\star)^T\mathbf{Z}(h;\theta^\star).$$

Δ_i is the unit square centered at i, $\mathbf{LPL}^{(1)}_{\Delta_i}(\varphi;\theta^\star)$ is the gradient vector of $LPL_{\Delta_i}(\varphi;\theta)$ (evaluated in $\theta^\star$), $\underline{\mathbf{U}}^{(2)}(\theta^\star)$ and $\mathbf{Z}(h;\theta^\star)$ are defined for $j,k=1,\dots,p$ by

$$(\underline{\mathbf{U}}^{(2)}(\theta^\star))_{j,k}=E(v_j(0^M|\Phi)v_k(0^M|\Phi)e^{-V\left(0^M|\Phi;\theta^\star\right)})$$

$$Z_j(h;\theta^\star):=E(v_j(0^M|\Phi)h(0^M|\Phi)e^{-V\left(0^M|\Phi;\theta^\star\right)}).$$

3 Applications to the empty space function

Definition 2 (Empty space function) *The empty space function F is the distribution of the distance from the origin (or another fixed point) to the nearest point in Φ, i.e.*

$$F(r)=P\left(\Phi\cap\mathcal{B}(0,r)\neq\emptyset\right)$$

It is known that the F function characterizes better the distribution of the spatial point process rather than one type of residuals. What is interesting in the approach developped in the previous section is that by taking the function

$$h_r(x^m|\varphi,\theta) := \mathbf{1}_{[0,r]}(d(x^m,\varphi)e^{V(x^m|\varphi;\theta)} = \tilde{h}(r)$$

then the residuals $R_{\Lambda_n}(\varphi;\tilde{h}(r))$ corresponds to a different of two estimations of the empty space function at distance r fixed (the first one is non parametric (since it equals to $|\Lambda_n|^{-1}\int_{\Lambda\times\mathbb{M}}\mathbf{1}_{[0,r]}(d(x^m,\varphi)\mu(dx^m)))$ while the second one is parametric).

Corollary 1 *Under the assumptions of Theorem 2 then for all* $(r_1,\ldots,r_d)^T \in (]0,+\infty[)^d$, *we have the following convergence in distribution as* $n\to+\infty$

$$|\Lambda_n|^{1/2}\ \Big(R_{\Lambda_n}(\Phi;\tilde{h}(r_1)),\ldots,R_{\Lambda_n}(\Phi;\tilde{h}(r_d))\Big)^T \to (\mathbb{G}(r_1;\theta^\star),\ldots,\mathbb{G}(r_d;\theta^\star))^T,$$

where $\mathbb{G}(\cdot;\theta^\star)$ *is a Gaussian process with covariance structure given, for* $r,s>0$, *by*

$$Cov\left(\mathbb{G}(r;\theta^\star),\mathbb{G}(s;\theta^\star)\right) = \sum_{i\in\mathbb{B}(0,\lceil D\rceil)} E\left(W_{\Delta_0}(\Phi;h(r),\theta^\star)W_{\Delta_i}(\Phi;h(s),\theta^\star)\right)$$

Hence, one can construct an asymptotic test based (for example) on the following result: choose $r_1,\ldots,r_d>0$ then the following convergence in distribution holds as $n\to+\infty$

$$|\Lambda_n|\sum_{i=1}^d R_{\Lambda_n}(\Phi;\tilde{h}(r_i))^2 \to \sum_{i=1}^d \mathbb{G}(r_i;\theta^\star)^2.$$

Note that in the different previous results, it is possible to tabulate the different distributions via Monte-Carlo experiments. A perspective of this work is to investigate the previous results in a simulation study and in practice. As a theoretical point of view, a perspective could be to obtain a weak convergence instead a finite-dimensional one.

References:

[1] A. Baddeley, R. Turner, J. Mller and M. Hazelton (2005) Residual analysis for spatial point processes, *Journal of the Royal Statistical Society*, 67(3), 617–666.

[2] A. Baddeley, J. Mller and A.G. Pakes (2008) Properties of residuals for spatial point processes, *Annals of the Institute of Statistical Mathematics*, 60(7), 627–649

[3] E. Bertin, J.-M. Billiot and R. Drouilhet (1999) k-nearest neighbour Gibbs point processes, *Markov processes and related fields*, 5(2), 219–234

[4] J.-M. Billiot, J.-F. Coeurjolly and R. Drouilhet (2008) Maximum pseudolikelihood estimator for exponential family models of marked Gibbs point processes, *Electronic Journal of Statistics*, 2, 234–264.