Estimation of a Regression Model and its Application in Renewal Process

Galaye Dia ¶ Abdoulaye Kone †

May 20, 2003

Abstract

We estimate a regression function on a point process by the Tukey regressogram method in a general setting and we give an application in the case of a cumulative Renewal Process.

AMS: primary 60F15;secondary 60G25
Keywords: Point process,regressogram,superposition,renewal

[1]

1 General hypotheses

We consider in this paper a bidimensional point process f_0 defined on a probability space $\left(\Omega, \mathcal{A}, P\right)$ with values in $\mathbf{R}_+ \times \mathbf{R}_+$. For any Borel set A of $\mathbf{R}_+ \times \mathbf{R}_+$ denote by $f_0(A)$ the number of points falling in A. We suppose that $l = f_0(\mathbf{R}_+ \times \mathbf{R}_+)$ is finite almost surely and that the mean measure of f_0 say μ admits a Radon Nikodym derivative f^*.For every open non-empty set Θ we suppose that $P(f_0(\Theta) > 0) > 0$.

Let $f_{0,1}$ be the first projection of f_0;we denote by μ_1 its mean measure and f its Radon Nikodym density. If $l \geq 1$ let (X_1, Y_1) ,...,(X_l, Y_l) the points of the process ordered such that $X_1 < ... < X_l$.
We define (X_0, Y_0)=$(0, 0)$.Let α be an integer and suppose $l = l_0$,$l_0 > 0$.

The model of regression we are considering satisfies the following:
a)$E\left(Y_j^\alpha / X_1 = x_1, ..., X_j = x_j, ..., X_{l_0} = x_{l_0}\right) = E\left(Y_j^\alpha / X_j = x_j\right)$ for $j = 1, ..., l_0$

b)$E\left(Y_j^\alpha / X_j = x\right)$ is independant of j and l_0 for $j = 1, ..., l_0$. We denote this funtion as $\Psi_\alpha(x)$

Consider f_i for $i = 1, ...n$ n i.i.d points processes having the same distribution as f_0 and $f_{(n)}$ their superposition in the sens of Cox[1966]. We denote $m = f_{(n)}(\mathbf{R}_+ \times \mathbf{R}_+)$ and $f_{1,(n)}$ the first projection of $f_{(n)}$. If α=1 we denote as Ψ the function Ψ_α.

The estimator we are dealing with is the fixed bandwidth regressogram of Tukey [1961] developped later by Major[1973], Geffroy[1980]. It was utilized for estimating the regression function on a Poisson Process in Dia[1987].

Suppose $m \geq 1$ and let $\left(X_1^{(n)}, Y_1^{(n)}\right)$,...,$\left(X_m^{(n)}, Y_m^{(n)}\right)$ be the points of $f_{(n)}$. If $m = 0$ we put $\left(X_0^{(n)}, Y_0^{(n)}\right) = (0, 0)$.

[1]¶mail to **UFR de Mathematiques Appliquees et d'Informatique**
B.P 234 Universite Gaston Berger de Saint-Louis (Senegal) E-mail: galaye@ugb.sn
&Laboratoire d Etudes et de Recherche en Statistiques et Developpement (LERSTAD)
Universite Gaston Berger de Saint-Louis (Senegal)
†Departement de Mathematiques Universite Cheick Anta Diop
Dakar (Senegal)

Let k be a function of n. We denote

$$\begin{aligned}
\Delta_{k,r} &= [\frac{r}{k}, \frac{r+1}{k}[, r \in N \\
\mathcal{J}_{n,r} &= \{i, i \geq 1\, X_i^{(n)} \in \Delta_{k,r}\} \\
\nu_{n,r} &= card\mathcal{J}_{n,r} \\
\overline{Y}_{n,r} &= \begin{cases} \frac{1}{\nu_{n,r}} \sum_{i\in\mathcal{J}_{n,r}} Y_i^{(n)} & \text{if } \nu_{n,r} > 0 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

We then define $\Psi_{n,k}$ the estimator of Ψ by
$(\forall r \geq 0) \quad (\forall x \in \Delta_{k,r}) \quad \Psi_{n,k}(x) = \overline{Y}_{n,r}$

2 The main theorems

Let $f_{0,1}$ be the first projection of f_0 and let us consider the following hypotheses:

- H_1) f is continuous and strictly positive
- H_2) Ψ_α exists and is continuous for $\alpha = 1, 2$
- H_3) to each x in $\mathbf{R}_+$
  $P\bigg(f_{0,1}\big([x, x + \Delta x[\big) \geq 2\bigg) = o(\Delta x)$
- H_4) $f_{0,1}$ satisfies the approximation (see Daley[1974])
  $P\bigg(f_{0,1}(I) = 1\bigg) \cong \mu(I)$
- H_5) the second factorial moment of $f_{0,1}(I)$ exists for every bounded interval I

Remark: It results of the hypothesis H_3 that $f_{0,1}$ is without double points, that is

$(\forall i, j)\ (1 \leq i <$j) $P\bigg(\varpi : X_i(\varpi) = X_j(\varpi) \quad l > 1\bigg)$=0
Therefore the points $X_1, ..., X_l$ can be strictily ordered with probability one (see Daley [1974]).

Theorem 1 *If, for $l = k \geq 1$ the expression $E(Y_j^\alpha / X_j = x)$ is finite and independant of j, $j = 1, ..., k$ and k then*

$$\Psi_\alpha(x) = \frac{1}{f(x)} \int_{\mathbf{R}} y^\alpha f^*(x, y)\, dy$$

where $f(x) = \int_{\mathbf{R}} f^*(x, y)\, dy$

Theorem 2 *Suppose the hypotheses H_1, H_2, H_3, H_4, H_5 are satisfied. If $k \to \infty$, $\frac{n}{k^2} \to \infty$ and $k = o(\frac{n}{Logn})$ as $n \to \infty$ then $\Psi_{n,k}$ is a weakly consistant estimator of Ψ i.e*

$$(\forall x \in \mathbf{R}) \quad \lim_{n\to\infty} E[(\Psi_{n,k}(x) - \Psi(x))^2] = 0$$

3 Preliminary Results

Lemma 3 *If $l = k$, then the variables $(X_1, Y_1), ..., (X_k, Y_k)$ are absolutely continuous with respect the Lebesgue measure with density $[P(l = k)]^{-1} h_k^i f^*$ $i = 1, ..., k$, say. Morover*

$$\sum_{k=1}^{\infty} (\sum_{i=1}^{k} h_k^i) = 1$$

Proof:The first part is a two-dimensional version of theorem 2.1 and its colloraly in Dia[1990] with $\phi = \chi_A$, A being any Borel set. That is: there exists Borel measurable function h_k^i such that

$$P((X_i, Y_i) \in A, l = k) = \int_A h_k^i(x, y)\, d\mu(x, y) = \int_A h_k^i(x, y) f^*(x, y)\, dx\, dy \qquad (1)$$

We have by that theorem

$$\sum_{k=1}^{\infty} \sum_{i=1}^{k} P((X_i, Y_i) \in A, l = k) = \mu(A) \qquad (2)$$

or $\sum_{k=1}^{\infty} (\int_A \sum_{i=1}^{k} h_k^i(x, y)\, dx\, dy) = \int_A f^*(x, y)\, dx\, dy$

The Beppo-Levi theorem implies that

$$\sum_{k=1}^{\infty} (\sum_{i=1}^{k} h_k^i(x, y) f^*(x, y)) = f^*(x, y)$$

The announced theorem is proved

A similar result is obtained for the one dimensional process $f_{0,1}$, say.The variables $X_i, i = 1, ..., k$ are absolutely continuous with respect to Lebesgue measure with conditional density $f_k^i(x) = [P(l = k)]^{-1} g_k^i(x) f(x)$ and $\sum_{k=1}^{\infty} [\sum_{i=1}^{k} g_k^i] = 1$

4 The proofs of the theorems

4.1 Theorem 1

From equality(1) and the hypothesis Ib) on the regression, we have

$$\Psi_\alpha(x) = E(Y_j^\alpha / X_j = x) = \frac{\sum_{k=j}^{+\infty} \int_{\mathbf{R}_+} y^\alpha h_k^j(x, y) f^*(x, y)\, dx\, dy}{\sum_{k=j}^{+\infty} g_k^j(x) f(x)} \qquad (3)$$

for each j. Which implies

$$\Psi_\alpha(x) = \frac{\sum_{k=1}^{+\infty} \sum_{j=1}^{k} \int_{\mathbf{R}_+} y^\alpha h_k^j(x, y) f^*(x, y)\, dx\, dy}{\sum_{k=1}^{+\infty} \sum_{j=1}^{k} g_k^j(x) f(x)} \qquad (4)$$

the serie in the numerator being convergent. It follows by the lemma 1 and the Beppo-Levi theorem the desired result.

Denote $r = [kx]$ For fixed x in $\mathbf{R}_+^*$ where $[z]$ stands for the entire part of a real number z. $\Delta_{k,r}^j = (\nu_{n,r} = j)$.

Consider the following partition of $\mathcal{J}_{n,r}$ say

$\mathcal{J}_{n,r} = \bigcup_{s=1}^n \mathcal{J}_{n,r}^{(s)}$ where $\mathcal{J}_{n,r}^{(s)}$ stands for the set of indexes i such that $X_i^{(n)}$, element of the s th component of $f_{1,(n)}$ denoted by $f_{1,s}$, belongs to $\Delta_{k,r}$. Let $\nu_{n,r}^{(s)} = card\mathcal{J}_{n,r}^{(s)}$

Lemma 4 *Suppose that H_1 H_2 H_3 are satisfied. Then there exist a point $\zeta_{k,r,\alpha}$ in the closure of $\Delta_{k,r}$ such that*
$(\forall j \geq 1) \quad (\forall i \in \mathcal{J}_{n,r}) \quad E((Y_i^{(n)})^\alpha / \Delta_{k,r}^j) = \Psi_\alpha(\zeta_{k,r,\alpha})$

Proof

$$\int_{\nu_{n,r}=j} (Y_i^{(n)})^\alpha \, dP = \sum_{\substack{j_1,j_2,\ldots j_s \ldots j_n \\ j_1+j_2\ldots+j_n=j}} \int_{\nu_{n,r}^{(s)}=j_s} (Y_i^{(n)})^\alpha \, dP$$

We make the convention that the integral in the right hand side is nul if $\mathcal{J}_{n,r}^{(s)} = \emptyset$ or if $i \notin \mathcal{J}_{n,r}^{(s)}$

Therefore if $\mathcal{J}_{n,r}^{(s)} \neq \emptyset$ and $i \in \mathcal{J}_{n,r}^{(s)}$ the hypothesis a) in I) gives

$$\int_{\nu_{n,r}=j_s, i \in \mathcal{J}_{n,r}^{(s)}} (Y_i^{(n)})^\alpha \, dP = \int_{X_i^{(n)} \in \Delta_{k,r}, \ldots X_{i_{j_s}}^{(n)} \in \Delta_{k,r}} \psi_\alpha(x_i) dF_{X_i^{(n)}, \ldots X_{i_{j_s}}^{(n)}}(x_i, \ldots, x_{i_{j_s}})$$

where $(X_i^{(n)}, \ldots, X_{i_{j_s}}^{(n)})$ are the j_s variables of the set $(\nu_{n,r}^{(s)} = j_s)$.

Since Ψ_α is continuous, we obtain

$$\int_{\nu_{n,r}^{(s)}=j_s} (Y_i^{(n)})^\alpha \, dP = P(\nu_{n,r}^{(s)} = j_s)\Psi_\alpha(\zeta_s)$$

ζ_s belonging to the closure of $\Delta_{k,r}$. Hence

$$\int_{\nu_{n,r}=j} (Y_i^{(n)})^\alpha \, dP = \sum_{\substack{j_1,j_2,\ldots j_s \ldots j_n \\ j_1+j_2\ldots+j_n=j}} P(\nu_{n,r}^{(s)} = j_s)\Psi_\alpha(\zeta_s)$$

$E\left((Y_i^{(n)})^\alpha / \Delta_{k,r}^j\right) = \frac{1}{P(\Delta_{k,r}^j)} \int_{\nu_{n,r}=j} (Y_i^{(n)})^\alpha \, dP$ is then between $min_{x \in \Delta_{k,r}} \Psi_\alpha(x)$ and $max_{x \in \Delta_{k,r}} \Psi_\alpha(x)$
Since Ψ_α is continuous, we get the announced result.

Proposition 1 *If, H_1, H_2,H_3, H_4,H_5 are satisfied and $\frac{n}{k^2} \to \infty$, $n \to \infty$ then*

- *a)* $\lim_{n \to +\infty} \sum_{j=1}^{+\infty} P(\Delta_{k,r}^j) = 1$
- *b)* $\lim_{n \to +\infty} \sum_{j=1}^{+\infty} \frac{1}{j} P(\Delta_{k,r}^j) = 0$
- *c)* $\lim_{n \to +\infty} n \sum_{j=1}^{+\infty} \frac{1}{j^2} P(\Delta_{k,r}^j) = 0$

Proof

Write

$$\sum_{j=1}^{\infty} P(\Delta_{k,r}^{j}) = 1 - P(\Delta_{k,r}^{0})$$

But

$$P(\Delta_{k,r}^{0}) = P(f_{0,1}(\Delta_{k,r}) = 0))^{n} = (1 - (P(f_{0,1}(\Delta_{k,r}) = 1) + P(f_{0,1}(\Delta_{k,r}) \geq 2))^{n}$$

and

$$\begin{aligned} P(f_{0,1}(\Delta_{k,r}) = 1) &= \mu_1(\Delta_{k,r}) + o(\tfrac{1}{k}) \\ &= \tfrac{1}{k} f(\alpha) + o(\tfrac{1}{k}) \end{aligned}$$

where $\alpha \in \Delta_{k,r}$ because of H_4 and the continuity of f

By $H3$ we have $P(f_{0,1}(\Delta_{k,r}) \geq 2) = o(\frac{1}{k})$

Hence

$$\begin{aligned} P(\Delta_{k,r}^{0}) &= e^{nLog(1-(\frac{1}{k}f(\alpha)+o(\frac{1}{k})))} \\ &\cong e^{\frac{-n}{k}(f(\alpha)+\epsilon(\frac{1}{k}))} \end{aligned}$$

Since $\frac{n}{k} \to \infty$ as $n \to +\infty$ and $f(\alpha) \to f(x) > 0$,a) is proved.

b) $\sum_{j=1}^{\infty} \frac{1}{j} P(\Delta_{k,r}^{j}) = E(\frac{1}{\nu_{n,r}})$

Let us show that $\nu_{n,r} \to +\infty$ with probability one. Let $A \geq 2$.
We have

$$P(f_{0,1}(\Delta_{k,r}) \leq A) \leq P(f_{0,1}(\Delta_{k,r}) = 1) + P(f_{0,1}(\Delta_{k,r}) \geq 2)$$

Hence by part a)and H_3 we get $P(f_{0,1}(\Delta_{k,r} \leq A) \to 0$ as $n \to +\infty$

It follows that it exists $\delta > 0$ such that $P(\nu_{n,r} > A) > \delta$. Therefore by the Borel-Cantelli lemma infinitely many events$(f_{1,s}(\Delta_{k,r}) > A)$ occur with probability one. Hence $\nu_{n,r} = \sum_{1}^{n} f_{1,s}(\Delta_{k,r}) \to +\infty$ with probability one.

The Lebesgue dominated convergence theorem implies the announced result.

c)It is the same to chow that $lim_{n \to +\infty} E(\dfrac{n}{\nu_{n,r}^{2}}) = 0$

Write

$$\frac{n}{\nu_{n,r}^2} = \frac{1}{n\mu_1^2(\Delta_{k,r})}(\frac{n\mu_1(\Delta_{k,r})}{\nu_{n,r}})^2 \tag{5}$$

$$\frac{n\mu_1(\Delta_{k,r})}{\nu_{n,r}} = \frac{n\mu_1(\Delta_{k,r})}{\sum_{s=1}^{n}\nu_{n,r}^{(s)}} \tag{6}$$

But $E(\frac{\nu_{n,r}^{(s)}}{\mu_1(\Delta_{k,r})}) = 1$ and the random variables $\frac{\nu_{n,r}^{(s)}}{\mu_1(\Delta_{k,r})}$ $s = 1, ..., n$ are independant and identically distributed. Hence the equality (6) tends to 1 with probability one by the strong low of large numbers and therefore it is bounded with probability one.

$$\mu_1(\Delta_{k,r}) = \int_{\frac{r}{k}}^{\frac{r+1}{k}} f(x)\,dx = \frac{f(\alpha)}{k} \tag{7}$$

with $\alpha \in \Delta_{k,r}$

$f(\alpha) \to f(x) > 0$ as $n \to +\infty$ by the continuity of f.

Therefore $\frac{n}{\nu_{n,r}^2} = O(\frac{k^2}{n})$ a.s

The Lebesgue dominated convergence theorem implies the desired result.

Remark:The strong law of large numbers as mentioned earlier is slightly different from the standard in the literature, because the distribution function of the variables $\frac{\nu_{n,r}^{(s)}}{\mu_1(\Delta_{k,r})}$ depends on n for $s = 1, ..., n$.To make it valid it is sufficient to show,following (e.g Billingsley,p.291), that $E\left(\frac{\nu_{n,r}^{(s)}}{\Delta_{k,r}}\chi_{[\frac{\nu_{n,r}^{(s)}}{\Delta_{k,r}} \geq n]}\right)$ tends to zero as $n \to +\infty$.Let Δ be a bounded interval containing $\Delta_{k,r}$ for n large,then this expectation is bounded by $E\Big(f_{0,1}(\Delta)\chi_{[f_{0,1}(\Delta) \geq n\mu_1(\Delta_{k,r})]}\Big)$. Since $n\mu_1(\Delta_{k,r})$ tends to $+\infty$ as $n \to +\infty$ we have the desired result.

Proposition 2 *Under the conditions of theorem 2, if* $H4$ *is satisfied then*

$$lim_{n \to +\infty} \sum_{j=1}^{+\infty} \frac{1}{j^2} \sum_{i \neq i'} cov(Y_{i'}^{(n)}, Y_i^{(n)}/\Delta_{k,r}^j) P(\Delta_{k,r}^j) = 0$$

Proof We suppose that i and i' belong to the same $\mathcal{J}_{n,r}^{(s)}$ and $card\mathcal{J}_{n,r}^{(s)} \geq 2$ otherwise the covarience is nul
Let s be fixed and i, i' belong to $\mathcal{J}_{n,r}^{(s)}$

The inequality

$$|cov(Y_{i'}^{(n)}, Y_i^{(n)}/\Delta_{k,r}^j)| \leq (E((Y_i^{(n)})^2/\Delta_{k,r}^j))^{\frac{1}{2}}(E((Y_{i'}^{(n)})^2/\Delta_{k,r}^j))^{\frac{1}{2}}$$

implies

$$\begin{aligned}
\sum_{i\neq i'} |cov(Y_{i'}^{(n)}, Y_i^{(n)}/\Delta_{k,r}^j)|P(\Delta_{k,r}^j) &\leq \sum_{i\neq i'} E(\chi_{\Delta_{k,r}^j}(E((Y_i^{(n)})^2/\Delta_{k,r}^j))^{\frac{1}{2}}(E((Y_{i\prime}^{(n)})^2/\Delta_{k,r}^j))^{\frac{1}{2}}) \\
&\leq \sum_{\alpha=2}^{j} E(\sum_{\substack{i\neq i' \\ \mathcal{J}_{n,r}^{(s)}}} \chi_{\Delta_{k,r}^j}\Psi_2(\zeta_{k,r,2})/card\mathcal{J}_{n,r}^{(s)} = \alpha)P(card\mathcal{J}_{n,r}^{(s)} = \alpha) \\
&\leq \sum_{\alpha=2}^{j} \alpha(\alpha-1)\Psi_2(\zeta_{k,r,2})P(\nu_{n-1,r} = j-\alpha)P(\nu_{n,r}^{(s)} = \alpha)
\end{aligned}$$

We have for such i and i'

$$\begin{aligned}
\sum_{j=1}^{+\infty}\frac{1}{j^2}\sum_{i\neq i'} |cov(Y_{i'}^{(n)}, Y_i^{(n)}/\Delta_{k,r}^j)|P(\Delta_{k,r}^j) &\leq \sum_{j=1}^{+\infty}\frac{1}{j^2}\sum_{\alpha=2}^{j} \alpha(\alpha-1)\Psi_2(\zeta_{k,r,2})P(\nu_{n-1,r} = j-\alpha)P(\nu_{n,r}^{(s)} = \\
&\leq \sum_{\alpha=1}^{+\infty}\alpha(\alpha-1)P(\nu_{n,r}^{(s)} = \alpha)\sum_{j=\alpha}^{+\infty}\frac{1}{j^2}P(\nu_{n-1,r} = j-\alpha) \\
&\leq \Psi_2(\zeta_{k,r,2})\eta_r^{(2)}\left(\sum_{j=1}^{+\infty}\frac{1}{j^2}P(\nu_{n-1,r} = j) + P(\nu_{n-1,r} = 0)\right)
\end{aligned}$$

where $\eta_r^{(2)}$ stands for the second factorial moment of $f_{0,1}(\Delta_{k,r})$

Hence for i and i' belonging to $\mathcal{J}_{n,r}$ we have

$$\sum_{j=1}^{+\infty}\frac{1}{j^2}\sum_{i\neq i'} |cov(Y_{i'}^{(n)}, Y_i^{(n)}/\Delta_{k,r}^j)|P(\Delta_{k,r}^j) \leq n\Psi_2(\zeta_{k,r,2})\eta_r^{(2)}\left(\sum_{j=1}^{+\infty}\frac{1}{j^2}P(\nu_{n-1,r} = j) + P(\nu_{n-1,r} = 0)\right)$$

In proposition 1 a) we have stated that

$$P(\Delta_{k,r}^0) = e^{nLog(1-(\frac{1}{k}f(\alpha)+o(\frac{1}{k})))} \cong e^{\frac{-n}{k}(f(\alpha)+\epsilon(\frac{1}{k}))}$$

It follows that $Log\Big(nP(\nu_{n-1,r})\Big) \cong Logn - \frac{(n-1)}{k}(f(\alpha)+\epsilon(\frac{1}{k}))$ which tends to $-\infty$ as $n \to +\infty$ The proposition 1 c) then implies the desired result.

4.2 Theorem 2

By Lemma 2 we have

$$E(\Psi_{n,k}(x)) = E(E(\Psi_{n,k}(x)/\nu_{n,r})) = \Psi(\zeta_{k,r,1})\sum_{j=1}^{+\infty}P(\Delta_{k,r}^j) \qquad (8)$$

In the same way

$$
\begin{aligned}
E(\Psi^2_{n,k}(x)) &= \sum_{j=1}^{+\infty} E(\Psi^2_{n,k}(x)/\Delta^j_{k,r})P(\Delta^j_{k,r}) \\
E(\Psi^2_{n,k}/\Delta^j_{k,r}) &= \frac{1}{j^2}\sum_i E((Y_i^{(n)})^2/\Delta^j_{k,r}) + \frac{1}{j^2}\sum_{i\neq i'} E(Y_{i'}^{(n)}Y_i^{(n)}/\Delta^j_{k,r})
\end{aligned}
$$

Express $\sum_{j=1}^{+\infty}\frac{1}{j^2}\sum_{i\neq i'} E(Y_{i'}^{(n)}Y_i^{(n)}/\Delta^j_{k,r})P(\Delta^j_{k,r})$ as

$$
\sum_{j=1}^{+\infty}\frac{1}{j^2}\sum_{i\neq i'} cov(Y_{i'}^{(n)},Y_i^{(n)}/\Delta^j_{k,r})P(\Delta^j_{k,r}) + \Psi^2(\zeta_{k,r,1})\sum_{j=1}^{+\infty}P(\Delta^j_{k,r}) - \Psi^2(\zeta_{k,r,1})\sum_{j=1}^{+\infty}\frac{1}{j}P(\Delta^j_{k,r})
$$

Proposition 1 and Proposition 2 imply that this last expression tends to $\Psi^2(x)$ as $n\to+\infty$

On the ather hand

$$
\sum_{1}^{+\infty}\frac{1}{j^2}\sum_i E((Y_i^{(n)})^2/\Delta^j_{k,r})P(\Delta^j_{k,r}) \leq \Psi_2(\zeta_{k,r,2})\sum_{j=1}^{+\infty}\frac{1}{j}P(\Delta^j_{k,r})
$$

The right hand side of this expression tends to 0 as $n\to+\infty$ by proposition 1. It follows that $E(\Psi^2_{n,k}(x))\to\Psi^2(x)$ as $n\to+\infty$

The proposition 2 again implies, by equality (8), that $E(\Psi_{n,k}(x))\to\Psi(x)$ as $n\to+\infty$. Hence $Var(\Psi_n(x))\to 0$ as $n\to+\infty$

Since $lim_{n\to+\infty}(Bias\Psi_{n,k}(x))^2 = 0$ the proof of theorem 2 is complete.

Remark

If there exists Y independant of the process such that $Y_i < Y$ for $i = 1,2,...$ then $E((Y_i^{(n)})^2/\Delta^j_{k,r}) < \Psi_1(\zeta_{k,r,1})E(Y)$ and the theorem remains valid if Y has a finite moment.Therefore,in this case, we shall restrict ourself to processes for which the general hypotheses and H_2 are verified for $\alpha = 1$.

5 Application

Let X_n $n = 1,2..$, be a sequence of i.i.d random variables having the same disribution as a variable X with values in $\mathbf{R}_+$

Put $S_n = X_1+X_2+...+X_n$, $F(x) = P(X<x)$ its distribution which we suppose continuous with density $\widehat{f}$ strictly positive and continuous. For every x in $\mathbf{R}_+$ define $N(x) = Sup(n : S_n\in[0,x])$.

We consider the process $f_{0,1}$ generated by S_n with mean measure μ_1 defined by

$$
\forall A\in\mathbf{R}_+\ \mu_1(A) = \sum_{n=1}^{+\infty}P(S_n\in A)
$$

which admits a continuous and strictly positive derivative. We denote $\mu_1([0,x]) = \mu_1(x) = \sum_{k=1}^{+\infty} F_k^*(x)$
where $F_k^*(x)$ stands for the n-convolution of the distribution F
It is well-known that hypothesis H_5 is satisfied.
Let us now show that $f_{0,1}$ satisfies also H_3 and H_4

1) for H_3
We have
$\{f_{0,1}([x, x+\Delta x]) \geq 2\} \subset \bigcup_{k=1}^{+\infty}(S_k \in [x, x+\Delta x], S_{k+1} \in [x, x+\Delta x])$

And

$$\begin{aligned}
P(S_k \in [x, x+\Delta x], S_{k+1} \in [x, x+\Delta x]) &= \int_x^{x+\Delta x} P(S_{k+1} \in [x, x+\Delta x]/S_k = u)\, dF_{S_k}(u) \\
&= \int_x^{x+\Delta x} P(X_{k+1} + u \in [x, x+\Delta x]/S_k = u), dF_{S_k}(u) \\
&= \int_x^{x+\Delta x} P(X_{k+1} + u \in [x, x+\Delta x])\, dF_{S_k}
\end{aligned}$$

because the variables X_k are independant

We now express the term in the integral as
$P(X_{k+1} + u \in [x, x+\Delta x]) = \int_{x-u}^{x+\Delta x-u} \widehat{f}(t)\, dt = \Delta x \widehat{f}(\zeta)$
where $\zeta \in [x-u, x+\Delta x - u]$

Hence

$$\begin{aligned}
P(f_{0,1}([x, x+\Delta x]) \geq 2) &\leq \sum_{k=1}^{+\infty} \Delta x \widehat{f}(\zeta) \int_x^{x+x\Delta x} dF_{S_k}(u) \\
&\leq \Delta x \widehat{f}(\zeta) \int_x^{x+\Delta x} d(\sum_{k=1}^{+\infty} F_{S_k}(u)) \\
&\leq \Delta x \widehat{f}(\zeta) \int_x^{x+\Delta x} d\mu_1(x) = \Delta x \widehat{f}(\zeta) \mu_1(\Delta x)
\end{aligned}$$

Since μ_1 is continuous we get

$$P(f_{0,1}([x, x+\Delta x]) \geq 2) = o(\Delta x)$$

2)for H_4 we have on the one hand

$$P(f_{0,1}([x, x+\Delta x]) = 1) = \sum_{n=0}^{+\infty} P(f_{1,0}([x, x+\Delta x]) = 1, N(x) = n) \tag{9}$$

But

$$P(f_{0,1}([x,x+\Delta x]) = 1/N(x) = n) \;=\; P(S_{n+1} - S_n \in [x, x+\Delta x]/S_n = x)$$
$$= \; P(X_{n+1} \in [x, x+\Delta x]$$

Thus we obtain from (9) the equality

$$P(f_{0,1}([x,x+\Delta x]) = 1) = \int_x^{x+\Delta x} dF(x) \tag{10}$$

On the other hand the renewal equation

$$\mu_1(t) = F(t) + \int_0^t \mu_1(t-u)\, dF(u)$$

gives

$$\mu_1([x,x+\Delta x]) = \int_x^{x+\Delta x} dF(x) + \int_x^{x+\Delta x} (\mu_1(x+\Delta x - u) - \mu_1(x-u))\, dF(u)$$
$$= \int_x^{x+\Delta x} dF(x) + \mu_1(\Delta x) \int_x^{x+\Delta x} dF(u) \tag{11}$$

The equalities (10) and (11) give the desired result.

Suppose now, with the ith variable X_i, is associated a second variable W_i such that the (X_i, W_i) are independant and W_i are identically distributed. We impose X_i and W_i to be dependant.

The process defined by

$$Y_{N_t} = \begin{cases} \sum_{i=1}^{N_t} W_i & \text{if } N_t > 0 \\ 0 & \text{otherwise} \end{cases}$$

is a cumulative process on $\mathbf{R}_+$
Suppose that $Z_1 = (X_1, W_1)$ have density $f(x,w)$, then the random vector $Z_2 = (X_1 + X_2, W_1 + W_2)$ admits a density given by

$$f^{(2)}(x,w) = \int_0^{+\infty} \int_0^{+\infty} f(x-u, w-v) f(u,v)\, du\, dv$$

where $f^{(2)}$ stands for the two-fold convolution of f (e.g. Pugachev 1965, p.128)

The process $Z_{N_t} = (\sum_{i=1}^{N_t} W_i, \sum_{i=1}^{N_t} X_i)$ admits mean measure μ with density f^* defined by

$$f^*(x,w) = \sum_1^{+\infty} f^{(n-1)} * f(x,w)$$

with $f^0 = 1$
We are now interested in the regression function $E(Y_r/S_r = x)$

Let f_{S_r} be the density of S_r and $f_{(S_r,X_2,W_2)}$that of (S_r, X_2, W_2)
We have

$$\begin{aligned} P(S_2 < x, X_2 < x_2, W_2 < w) &= \int_0^{x_2}\int_0^{w} P(S_2 < x/X_2 = u, W_2 = v) dF_{(X_2,W_2)}(u, v) \\ &= \int_0^{x_2}\int_0^{w} P(X_1 < x - u) f(u, v)\, du\, dv \end{aligned}$$

We deduce that

$$f_{(S_2,X_2,W_2)}(x, x_2, w) = f_{X_1}(x - u) f(x_2, w)$$

Hence

$$f_{(S_2,W_2)}(x, w) = \int_0^{x} f_{X_1}(x - u) f(u, w)\, du$$

A similar calculus leads to the following result

$$f_{(S_r,W_r)}(x, w) = \int_0^{x} f_{S_r}(x - u) f(u, w)\, du$$

We get from this:

$$E(W_r/S_r = x) = \frac{1}{f_{S_r}(x)} \int_0^{+\infty} w f_{(S_r,W_r)}(x, w)\, dw$$

$$= \frac{1}{f_{S_r}(x)} \int_0^{x} [f_{S_{r-1}}(x - u) \int_0^{+\infty} w f(u, w)\, dw]\, du$$

$$= \frac{1}{f_{S_r}(x)} \int_0^{x} f_{S_{r-1}}(x - u) f_{X_r}(u) E(W_r/X_r = u)\, du \tag{12}$$

hence

$$E(Y_r/S_r = x) = \frac{r}{f_{S_r}(x)} \int_0^{x} f_{S_{r-1}}(x - u) f_{X_r}(u) E(W_r/X_r = u)\, du$$

We are now in position to examine the conditions under which this regression function is independant of the index r.
If it is so we have:

$$\Psi(x) = \int_0^{x} \frac{\sum_{r=2}^{+\infty} r f_{S_r}(x - u) f_{X_1}(u) \Psi(u)}{\sum_{r=2}^{+\infty} f_{S_r}(x)}\, du \tag{13}$$

Puting

$$K(x,u) = \frac{\sum_{r=2}^{+\infty} r f_{S_r}(x-u) f_{X_1}(u)}{\sum_{r=2}^{+\infty} f_{S_r}(x)}$$

We recognize in (13) a Voltera equation of the the second kind with kernel $K(x,u)$.

Suppose that X is exponentially distributed with density $\rho e^{-\rho x}$ then N_t is a Poisson process and S_r has a Gamma distribution with density $\rho(\rho x)^{r-1}e^{-\rho x}/(r-1)!$. Straightforward calculations give:

$$K(x,u) = \frac{\rho(\rho(x-u)+2)e^{\rho(x-u)}}{e^{\rho x}-1}$$

Theorem 5 *Suppose that* $\Psi(x)$ *admits a continuous derivative and* $\Psi(0)=0$. *Then* $E(Y_r/S_r = x)$*is independant of* r *if and only if* $E(W_1/X_1 = x)$ *is linear.*

Proof

Suppose that $E(Y_r/S_r = x) = \Psi(x)$ independant of r. We can write,using (12)

$$\Psi(x) = \frac{r\int_0^x (x-u)^{r-1}\Psi'\,du}{x^{r-1}} \tag{14}$$

We get the differential equation

$$\Psi' x^{r-1} + (r-1)\Psi x^{r-2} = r\int_0^x (r-1)(x-u)^{r-2}\Psi'(u)\,du$$

The integral in the second member is equal to Ψx^{r-2}

so we obtain the equation

$$\Psi' x - \Psi = 0$$

which leads to the necessary condition. It is evident that if the derivative of $E(W_1/X_1 = x) = \lambda x$ is replaced in the integral(14) we obtain the reciprocal of the theorem.

Remark

The necessity of the theorem can be obtained with the Voltera equation.By successive integration by parts we get

$$\int_0^x \Psi'(t)e^{-\rho t}\,dt = \Psi'(x)(1-e^{-\rho x})/\rho$$

Letting ρ tends to zero we obtain the desired result.

Let us now verify hypothesis Ib) for $\alpha = 1$. We have for $i = 1, ..., r$

$$E(W_i/S_1 = x_1, S_2 = x_2, ..., S_i = x_i, ...S_r = x_r)$$
$$= \int_0^{+\infty} \frac{w f_{(W_i, X_1, X_2, ..., X_i, ..., X_r)}(w, x_1, x_2 - x_1, ..., x_i - x_{i-1}, ..., x_r - x_{r-1})}{f_{(X_1, X_2, ..., X_i, ...X_r)}(x_1, x_2 - x_1, ..., x_i - x_{i-1}, ..., x_r - x_{r-1})} dw$$

Using the independance of the variables we get

$$E(W_i/S_1 = x_1, S_2 = x_2, ..., S_i = x_i, ...S_r = x_r)$$
$$= \int_0^{+\infty} \frac{w f_{(X_i, W_i)}(x_i - x_{i-1}, w) f_{(X_1, ..., X_{i-1}, X_{i+1}, ..., X_r)}(x_1, ..., x_{i-1} - x_{i-2}, x_{i+1} - x_i, ..., x_r - x_{r-1})}{f_{X_i}(x_i - x_{i-1}) f_{(X_1, ..., X_{i-1}, X_{i+1}, ..., X_r)}(x_1, ...x_{i-1} - x_{i-2}, x_{i+1} - x_i, ...x_r))} dw$$
$$= \int_0^{+\infty} \frac{w f_{(X_i, W_i)}(x_i - x_{i-1}, w)}{f_{X_i}(x_i - x_{i-1})} dw$$
$$= \lambda(x_i - x_{i-1})$$

Puting $x_0 = 0$, it follows that

$$E(Y_k/S_1 = x_1, S_2 = x_2, ..., S_k = x_k, ...S_r = x_r) = E(Y_k/S_k = x_k) \text{ for } k = 1, ..., r.$$

We suppose the existence of a random variable Y_t independant of the proceess such that $Y_{N_t} < Y_t$ and $E(Y_t) < +\infty$ for each t then, with the remark following theorem 2,the conditions are satisfied to apply the method of estimation on arbitrarily interval $[0, T]$ of $\mathbf{R}_+$.

References

[1] **Billingsley ,P**(1986)*Probability and Measure* (Wiley,New York, 2nd ed.)

[2] **Cox ,D.R**(1966)*Theorie du Renouvellement* (Dunod, Paris)

[3] **Daley,D.J**(1974) Various concepts of orderliness for point processes. in;E.F. Harding and D.G Kendall. *Stochastic Geometry* (Wiley,New York)

[4] **Daley,D.J Vere-Jones,D**(1988) *An Introduction to the Theory of Point Processes* (Springer-Verlag)

[5] **Dia,G**(1987) Estimation of a Regression function on a Poisson process *Stat.Prob.Letters* N^o1,47-54

[6] **Dia,G**(1990) Non-parmetric Estimation of the density of a Point Process *Stat.Prob.Letters* N^o6,397-405

[7] **Geffroy,J**(1980) Etude de la convergence du regressogramme *Publ.Inst.Stat.univ.Paris* **xxv**,1-2,41-56

[8] **Karling,S Taylor H.M**(1990) *A first Course in Stochastic Processes*(Academic Press,2nd.ed)

[9] **Major,P**(1973) On a non parametric estimation of the regression function, *Studia.Sc.Math.Hung* 8 ;347-361

[10] **Pugachev ,V.S**(1965) *Theory of random function and its application to control problems* (Pergamon Press)

[11] **Tukey,J.W**(1961) Curves as parameters and touch estimation, *Proc.fourth.Berckley. Sympos. Math.Stat.Prob. Vol1* 681-694