MEAN-SQUARE CONVERGENCE OF A KERNEL-TYPE ESTIMATE OF THE INTENSITY FUNCTION OF AN INHOMOGENEOUS POISSON PROCESS

Anna Kitayeva

Tomsk Polytechnic University, Tomsk, Russia, *kit1157@yandex.ru*

Poisson processes are known to be useful to model several random phenomena (see for instance [1]). There are a lot of papers devoted to the problem of the intensity function estimation at a given point both under parametric and non parametric assumptions [2, 3].In the present paper we do not assume any parametric form of the function except that it is continuous at the point, and suppose that only a single realization of the process is available on $[0,T]$.

Let $\{t_i,\ \ i=\overline{1,N},\ \ 0\le t_i\le T\}$ be a realization of a Poisson point process having unknown intensity function $\lambda(t)$ on some time interval $[0,T]$, where *N* is the number of points falling into $[0,T]$.

It is well known that the distribution of *N*

$p_n = P(N=n) = \frac{\Lambda(0,T)^n}{n!}e^{-\Lambda(0,T)},\ n\ge 0,\ \Lambda(0,T)=\int_0^T \lambda(t)dt$, and conditionally to the event "the number of points *N* falling into $[0,T]$ is fixed", the points of the process $\{t_i\}$ obey the same law with distribution density function $\lambda(t)/\Lambda(0,T)$.

It is natural to take the following expression as an estimate of the function $\lambda(t)/\Lambda(0,T)$ at a point *t*

$$S = \frac{1}{Nh_N}\sum_{i=1}^{N} K\left(\frac{t-t_i}{h_N}\right), \qquad (1)$$

where (h_n) is a sequence of positive real numbers such that $h_n \downarrow 0$ and $nh_n \to \infty$; the kernel $K(\cdot)$ is a compact real valued Borel function on $[-T,T]$ such that $\int_{-T}^{T} K(u)du = 1$.

Note that a sample size is a random quantity and we deal with the estimate based on random number of observations. Such kinds of kernel estimates were studied under some restrictions in [4].

Joint distribution of t_i and *N* [5]

$$p_{in}(x) = \lim_{\Delta x\to 0}\frac{P(t_i < x+\Delta x, N=n) - P(t_i < x, N=n)}{\Delta x} = \frac{\Lambda(0,x)^{i-1}}{(i-1)!}\frac{\Lambda(x,T)^{n-i}}{(n-i)!}e^{-\Lambda(0,T)}\lambda(x),$$

$0 < x < T,\quad n \ge i \ge 1.$

Consider an asymptotic behavior of statistic (1) under following scheme of series: let series of observations are done on $[0,T]$ with the intensity of the process in *n*-th trial equals to $\lambda_n(t) = n\lambda(t)$. Denote the value of the statistic (1) in *n*-th trial

$$S_n = \frac{1}{N_n h_{N_n}}\sum_{i=1}^{N_n} K\left(\frac{t-t_{in}}{h_{N_n}}\right), \qquad (2)$$

where N_n and (t_{in}) – respectively the number of observations and the realization of the process in *n*-th trial.

Theorem 1 (asymptotic unbiasedness). Let the kernel $K(\cdot)$ and the intensity function $\lambda(\cdot)$, in addition, satisfy the following conditions $\int_{-T}^{T}|K(u)|du<\infty,\ \sup_{x\in[0,T]}\lambda(x)<\infty,\ \lambda(\cdot)$ – continuous function at the point *t*. Then statistic (1) is asymptotically ($n\to\infty$ in (2)) unbiased.

Proof. Let $p_i(x/n)$ be the conditional density function t_i given $N=n$, then the expected value

$$E(S)=\sum_{n=1}^{\infty}p_n\frac{1}{nh_n}\sum_{i=1}^{n}\int_0^T K\left(\frac{t-x}{h_n}\right)p_i(x/n)dx=$$

$$=\sum_{n=1}^{\infty}\frac{1}{nh_n}\sum_{i=1}^{n}\int_0^T K\left(\frac{t-x}{h_n}\right)\frac{\Lambda(0,x)^{i-1}}{(i-1)!}\frac{\Lambda(x,T)^{n-i}}{(n-i)!}e^{-\Lambda(0,T)}\lambda(x)dx.$$

Taking into account $\sum_{i=1}^{n}\frac{\Lambda(0,x)^{i-1}}{(i-1)!}\frac{\Lambda(x,T)^{n-i}}{(n-i)!}=\frac{\Lambda(0,T)^{n-1}}{(n-1)!}$ we obtain

$$E(S)=\sum_{n=1}^{\infty}\frac{\Lambda(0,T)^{n-1}}{n!h_n}e^{-\Lambda(0,T)}\int_0^T K\left(\frac{t-x}{h_n}\right)\lambda(x)dx.$$

Denote $I_n=\frac{1}{h_n}\int_0^T K\left(\frac{t-x}{h_n}\right)\lambda(x)dx=\int_{(t-T)/h_n}^{t/h_n}K(x)\lambda(t-h_nx)dx$, $\Delta_n=I_n-\lambda(t)$.

Let us consider

$$\begin{aligned}|E(S_n)-\lambda(t)/\Lambda(0,T)|&=\frac{1}{\Lambda(0,T)}\left|\sum_{k=0}^{\infty}\frac{\Lambda(0,T)^k n^k}{k!}\Delta_k e^{-n\Lambda(0,T)}\right|=\\&=\frac{1}{\Lambda(0,T)}\left|\sum_{k=0}^{K}\frac{\Lambda(0,T)^k n^k}{k!}\Delta_k e^{-n\Lambda(0,T)}+\sum_{k=K+1}^{\infty}\frac{\Lambda(0,T)^k n^k}{k!}\Delta_k e^{-n\Lambda(0,T)}\right|\le\\&\le Cn^K e^{-n\Lambda(0,T)}+\sup_{k>K}|\Delta_k|\frac{1}{\Lambda(0,T)}\sum_{k=0}^{\infty}\frac{\Lambda(0,T)^k n^k}{k!}e^{-n\Lambda(0,T)}=\\&=Cn^K e^{-n\Lambda(0,T)}+\sup_{k>K}|\Delta_k|\frac{1}{\Lambda(0,T)},\end{aligned}\tag{3}$$

where $\Delta_0=0, C$ – some constant in respect of *n*.

Take arbitrary $\varepsilon>0$. If $h_n<\min(t/T,(T-t)/T)$, then from compactness and normalization $K(\cdot)$ it follows

$$|\Delta_n|=\left|\int_{-T}^{T}K(x)(\lambda(t-h_nx)-\lambda(t))dx\right|\le\sup_{x\in[-T,T]}|\lambda(t-h_nx)-\lambda(t)|\int_{-T}^{T}|K(u)|du.$$

Therefore inequality $|\Delta_n|\le\frac{\varepsilon}{2}\Lambda(0,T)$ holds for all sufficiently large *n*. Thus, for sufficiently large K $\sup_{k>K}|\Delta_k|\frac{1}{\Lambda(0,T)}\le\varepsilon,$ and the first term on the right-hand side of (3) obviously tends to zero as $n\to\infty$. The theorem is proved.

Joint distribution of t_i,t_j,N [5]

$$p_{ijn}(x,y)=\lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}}\frac{P(t_i<x+\Delta x,t_j<y+\Delta y,N=n)-P(t_i<x,t_j<y,N=n)}{\Delta x\Delta y}=$$

$$=\frac{\Lambda(0,x)^{i-1}}{(i-1)!}\frac{\Lambda(x,y)^{j-i-1}}{(j-i-1)!}\frac{\Lambda(y,T)^{n-j}}{(n-j)!}e^{-\Lambda(0,T)}\lambda(x)\lambda(y),$$

$$0<x<y<T,1\le i<j\le n,n\ge 2.$$

Theorem 2 (mean-square convergence). Let the assumptions of Theorem 1 hold and $\int_{-T}^{T}K^2(x)dx<\infty$. Then $\lim_{n\to\infty}E\left(S_n-\lambda(t)/\Lambda(T)\right)^2=0.$

Proof. The mean-square error

$\mathrm{MSE}(S_n)=E(S_n-\lambda(t)/\Lambda(T))^2=E(S_n-E(S_n))^2+(E(S_n)-\lambda(t)/\Lambda(T))^2=Var(S_n)+b(S_n)^2$, where the first term is called the variance and the second one is called the bias**.** As shown in Theorem 1 the bias $b(S_n)\underset{n\to\infty}{\to}0$. The variance

$$Var(S_n)=E(S_n{}^2)-(E(S_n))^2=$$

$$=E\left\{\frac{1}{(N_nh_{N_n})^2}\left[\sum_{i=1}^{N_n}K^2\left(\frac{t-t_{in}}{h_{N_n}}\right)+\sum_{i>j}K\left(\frac{t-t_{in}}{h_{N_n}}\right)K\left(\frac{t-t_{jn}}{h_{N_n}}\right)+\sum_{i<j}K\left(\frac{t-t_{in}}{h_{N_n}}\right)K\left(\frac{t-t_{jn}}{h_{N_n}}\right)\right]\right\}-$$

$$-\left\{E\left[\frac{1}{N_nh_{N_n}}\sum_{i=1}^{N_n}K\left(\frac{t-t_{in}}{h_{N_n}}\right)\right]\right\}^2=\frac{1}{\Lambda(0,T)}\sum_{k=1}^{\infty}\frac{1}{kh_k}\int_{(t-T)/h_{k_n}}^{t/h_{k_n}}K^2(x)\lambda(t-h_kx)dx\frac{\Lambda(0,T)^kn^k}{k!}\times$$

$$\times e^{-n\Lambda(0,T)}+2\sum_{k=2}^{\infty}\frac{1}{k^2h_k{}^2}\int_0^TK\left(\frac{t-x}{h_k}\right)\lambda(x)\int_x^TK\left(\frac{t-y}{h_k}\right)\lambda(y)dydx\times \tag{4}$$

$$\times\sum_{i=1}^{k-1}\sum_{j=i+1}^{k}\frac{\Lambda(0,x)^{i-1}}{(i-1)!}\frac{\Lambda(x,y)^{j-i-1}}{(j-i-1)!}\frac{\Lambda(y,T)^{k-j}}{(k-j)!}n^ke^{-n\Lambda(0,T)}-\frac{1}{\Lambda(0,T)^2}\left[\sum_{k=1}^{\infty}\frac{\Lambda(0,T)^kn^k}{k!}I_ke^{-n\Lambda(0,T)}\right]^2.$$

Note that according to Newton's polynom

$$\sum_{i=1}^{k-1}\sum_{j=i+1}^{k}\frac{\Lambda(0,x)^{i-1}}{(i-1)!}\frac{\Lambda(x,y)^{j-i-1}}{(j-i-1)!}\frac{\Lambda(y,T)^{k-j}}{(k-j)!}=\frac{\Lambda(0,T)^{k-2}}{(k-2)!}.$$

Denote $J_k=\frac{1}{h_k{}^2}\int_0^TK\left(\frac{t-x}{h_k}\right)\lambda(x)\int_x^TK\left(\frac{t-y}{h_k}\right)\lambda(y)dydx=$

$$=\int_{(t-T)/h_k}^{t/h_k}K(x)\lambda(t-h_kx)\int_{(t-T)/h_k}^{x}K(y)\lambda(t-h_ky)dydx.$$

Under the assumptions of Theorem 1 $\lim_{h_k\to 0}J_k=\lambda^2(t)\int_{-K}^{K}K(x)\int_{-K}^{x}K(y)dydx=\frac{\lambda^2(t)}{2}.$

Thus, the second term on the right-hand side of (4)

$$\frac{2}{\Lambda(0,T)^2}\sum_{k=2}^{\infty}\frac{k-1}{k}\frac{\Lambda(0,T)^kn^k}{k!}J_ke^{-n\Lambda(0,T)}\to\frac{\lambda^2(t)}{\Lambda(0,T)^2}\text{ as }n\to\infty.$$

Denote $L_k=\int_{(t-T)/h_k}^{t/h_k}K^2(x)\lambda(t-h_kx)dx.$ Taking into account

$\lim_{h_k \to 0} L_n = \lambda(t) \int_{-T}^{T} K^2(x)dx < \infty$ and $\lim_{n\to\infty} \frac{1}{nh_n} = 0$ we have that

$\lim_{n\to\infty} Var(S_n) = \frac{1}{\Lambda(0,T)} \lim_{n\to\infty} \sum_{k=1}^{\infty} \frac{1}{kh_k} L_k \frac{\Lambda(0,T)^k n^k}{k!} e^{-n\Lambda(0,T)} = 0.$ The theorem is proved.

As to choice of the bandwidth in (1) the method of cross-validation developed by M. Rudemo [6] and M.M. Brooks, J.S. Marron [7] for kernel estimators can be recommended.

Literature

1. J.F.C. Kingman. Poisson processes. Oxford Studies in Probability (1993) 112 p.
2. Yu.A Kutoyants. Statistical Inference for Spatial Poisson Processes. Lecture Notes in Statistics. v. 134, Springer, New York (1998) 276 p.
3. R. Helmers, I.W. Mangku, R. Zitikis. Statistical properties of a kernel type estimator of the intensity function of a cyclic Poisson process. J. Multivariate Anal, v. 92, (2005) pp. 1–23.
4. A.A. Nazarov, A.F. Terpugov. Queueing Theory. (In Russian) NTL Publishing House, Tomsk (2004) 228 p.
5. R.C. Srivastava. Estimation of Probability Density Function based on Random Number of Observations with Applications. Int. Stat. Rev., v. 41(1), (1973) p. 77-86.
6. M. Rudemo. Empirical choice of histograms and kernel density estimators. Scand. J. Stat. theory Appl., 9, (1982) pp. 65–78.
7. M.M. Brooks, J.S. Marron. Asymptotic optimality of the least-squares cross-validation bandwidth for kernel estimates of intensity functions. Stochastic Process Appl., v. 38(1), (1991) pp. 157–165.