BAYESIAN NONPARAMETRIC ESTIMATION OF POISSON INTENSITY ON A REAL LINE

Jana Timková

timkova@karlin.mff.cuni.cz
Department of Statistics, Charles University, Prague

SUMMARY

This poster presents a way of estimating an intensity of nonhomogeneous Poisson point process on a real line based on techniques introduced by Arjas, E., Gasbarra, D. (1994). The main idea is utilizing a piecewise constant function with varying number and size of intervals. The accuracy of the fit is assessed using the residual analysis.

Model of intensity

Let us suppose **a piecewise constant model for a Poisson process intensity function** $\lambda(t)$ during a time interval $(0, T]$. The intensity function is assumed to be constant within $(m+1)$ intervals which come from dividing $(0, T]$ by m random jump times $T_1, \ldots, T_m$. The level of $\lambda(t)$ regarding the interval $(T_{j-1}, T_j]$ is denoted as λ_j.

The number of intervals varies through **adding a new jump or deleting an existing one**.

Prior distribution with $\alpha_0, \beta_0, a_\mu, b_\mu, a_\alpha, b_\alpha$ being hyperparameters is set as:

- m jump times $T_1, \ldots, T_m \sim$ Poisson process with intensity μ and
- levels $\lambda_1, \ldots, \lambda_{m+1}$ with gamma priors
  $\lambda_1 \sim \Gamma(\alpha_0, \beta_0)$
  $\lambda_j \sim \Gamma(\alpha, \alpha/\lambda_{j-1}), \quad j = 2, \ldots, m+1$
- μ and α with gamma priors
  $\mu \sim \Gamma(a_\mu, b_\mu)$
  $\alpha \sim \Gamma(a_\alpha, b_\alpha)$

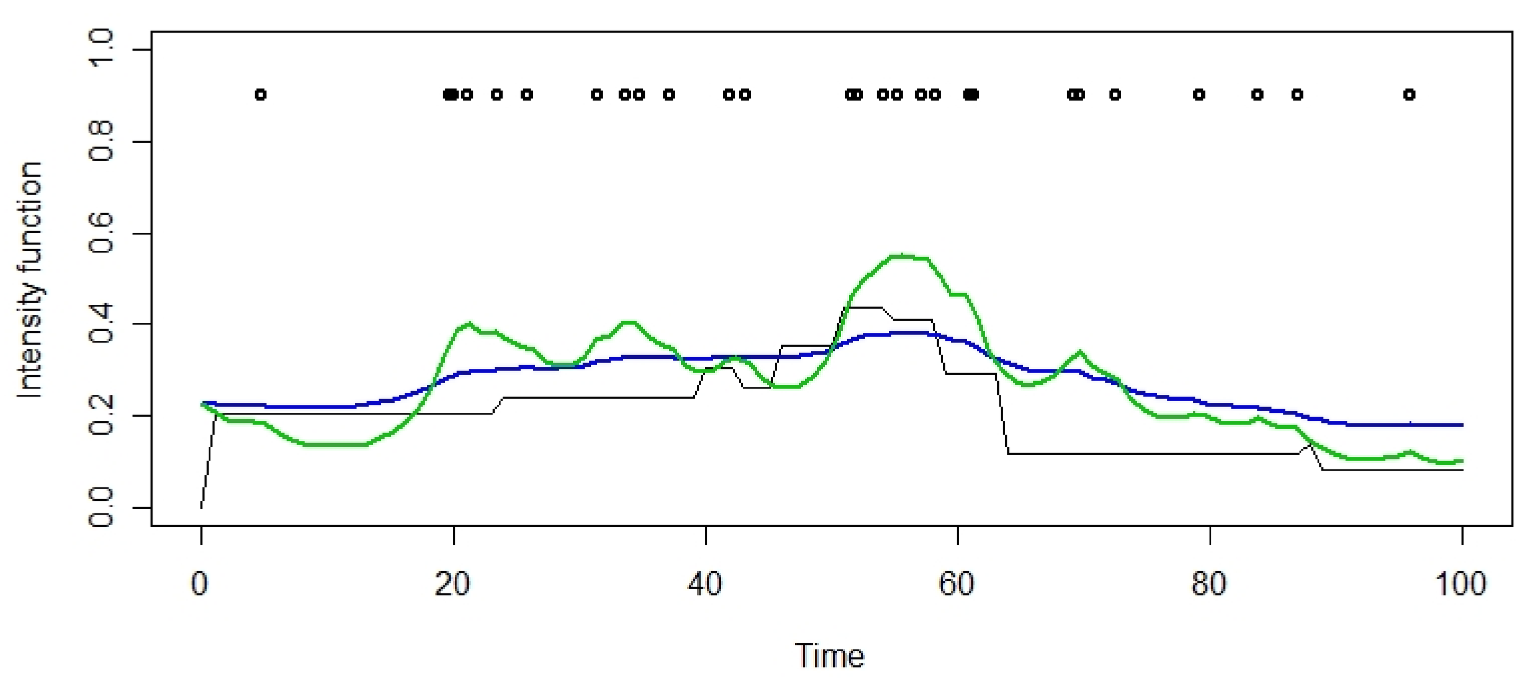
Posterior Distribution can be specified from the prior distribution of the intensity function and the likelihood of concrete realization $\mathbf{t} = (t_1, t_2, \ldots, t_n)'$, which is proportional to

$$L(\mathbf{t}) \propto \prod_{i=1}^{n} \lambda(t_i) \exp\Big\{ -\int_0^T \lambda(t)\, dt \Big\}.$$

To obtain the estimated intensity $\hat{\lambda}(t)$ we use the hybrid Gibbs sampler.

Simulated Data

A realization of a Poisson point process is simulated from a piecewise constant intensity. The picture displays the true intensity function (black line), the realization and two estimations of the intensity, with various choice of hyperparameters ($\alpha_0 = 1, \beta_0 = 20$ in both cases, blue line: $a_\mu = 1, b_\mu = 50, a_\alpha = 100, b_\alpha = 100$, green line: $a_\mu = 10, b_\mu = 20, a_\alpha = 10, b_\alpha = 10$).

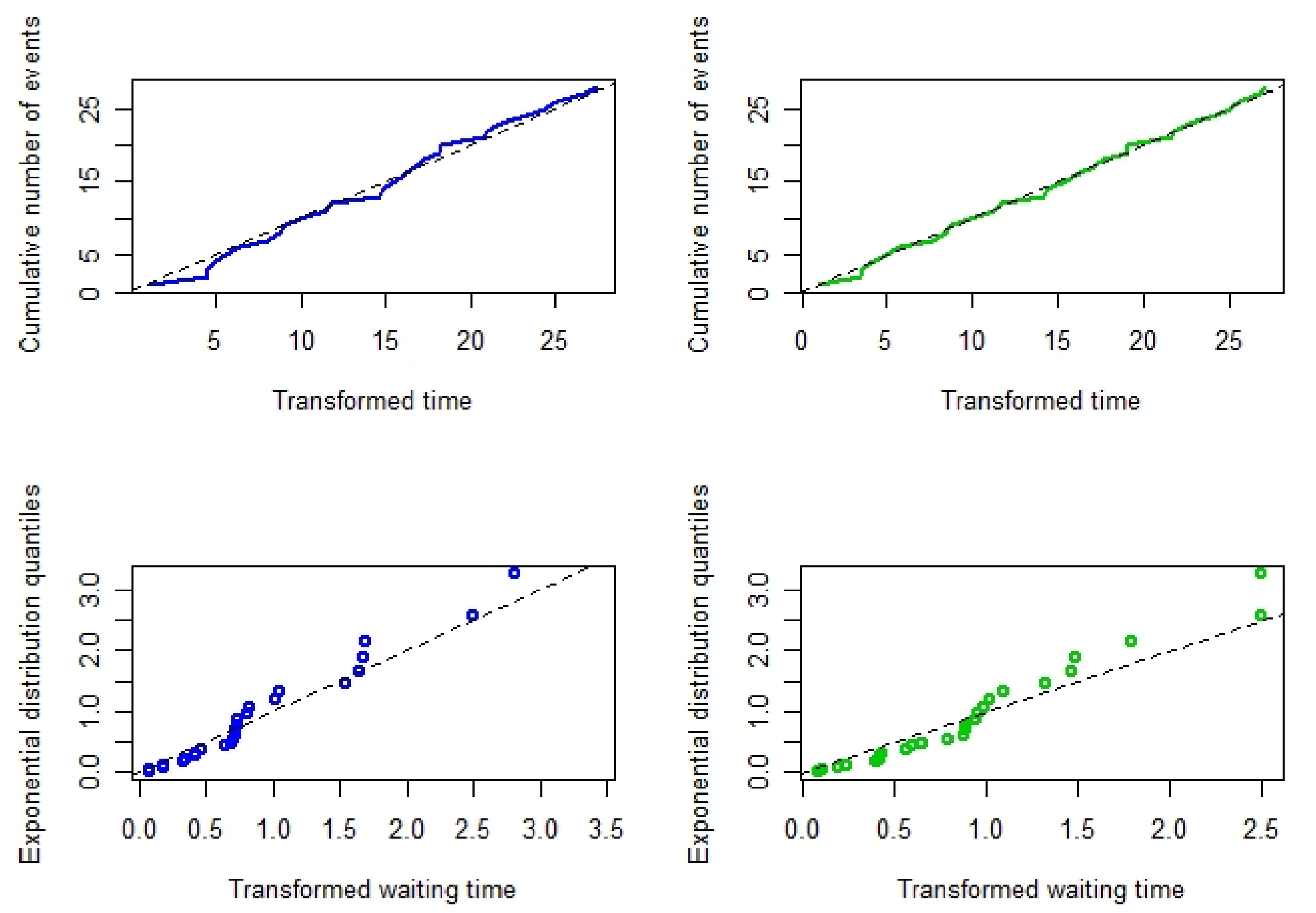


Residual Analysis

Considering the cumulative intensity function $\Lambda(t) = \int_0^t \lambda(s)\, ds$ and rescaling the time from t_i to $\tau_i = \Lambda(t_i)$ we get $\{\tau_i\}$ with the distribution of **a stationary unit intensity Poisson process**.

Hence if the estimated conditional intensity $\hat{\lambda}(t)$ is a good approximation to the true $\lambda(t)$, the transformed random times are expected to behave like an unit intensity Poisson process. In practise the integral could be approximated by MCMC simulations.

In the following figure the cumulative number of events is plotted against the transformed time. Futher the transformed waiting times $\{\tau_{i+1} - \tau_i\}$ distribution is assessed by Q-Q plots against the theoretical quantiles of the exponential distribution.



Outline

The presented method of estimating the intensity function of the Poisson point process in time should be viewed as a simple implementation of the Bayesian modeling, as it allows us to **switch between subspaces with various dimensions** represented by models with various number of parameters.

In **my future work** more emphasis will by put on

- nonparametric Bayesian estimation of an intensity function in spatial point processes using Voronoi's tessellation
- investigation into the Bayesian approach to residual analysis of point processes which takes the uncertainty about unknown parameters into account

Acknowledgement. The poster was supported by grants GA ČR 201/05/H007 and by GA AV IAA101120604.

References.

[1] Arjas, E., Gasbarra, D. (1994): Nonparametric bayesian inference from right censored survival data, using Gibbs sampler, *Statistica Sinica* 4, pp. 505-524.

[2] Green, P. J. (1995): Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, *Biometrika* 82, pp. 711-732.

[3] Ogata, Y. (1988): Statistical models for earthquake occurrences and residual analysis for point processes, *Journal of the American Statistical Association*, Vol. 83, No. 401, pp. 9-27