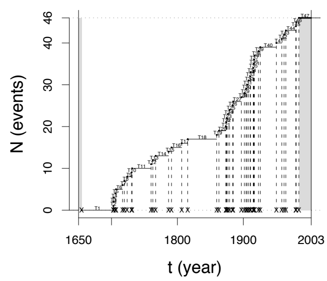
On Bayesian Nonparametric Estimation of Smooth Hazard Rates with a View to Seismic Hazard Assessment

Luca La Rocca
University of Modena and Reggio Emilia
larocca.luca@unimore.it

Research Report 38-05
http://www.cei.unimore.it

Seismic hazard assessment (counting process framework)



$N(t)$ counts the "strong" earthquakes up to time t in a given (Italian) seismogenic zone

The **geophysical risk** λ is the instantaneous conditional expected number of events per time unit (formally, the stochastic intensity of N with respect to its observed history)

Assuming **exchangeable inter-event times** $T_1, T_2, \ldots$ is not uncommon, usually in combination with of a parametric model; this gives

$$\lambda(t) = \widehat{\rho}\left(t - S_{N(t)}\right)$$

where S_i is the time of the i-th event and $\widehat{\rho}$ is the posterior pointwise expected hazard rate of the unknown inter-event time distribution

The **nonparametric point of view** has the advantage of giving a time-varying (possibly non-monotone) geophysical risk assessment without imposing any functional form on λ

Prior hazard rate proposal

$$\rho(t) = \xi_0 k_0(t) + \sum_{j=1}^{\infty} \xi_j k(t - \sigma_j), \quad t \geq 0$$

- $\xi_0, \xi_1, \xi_2, \ldots$ are $i.i.d.$ and *positive*
- $\sigma_j = \tau_1 + \cdots + \tau_j$ for $j \geq 1$
- $\tau_1, \tau_2, \ldots$ are $i.i.d.$ with *exponential* law
- ξ and τ are independent
- k is a probability density on $\mathbb{R}$
- k_0 is a positive function on $\mathbb{R}_+$ which is integrable in a neighbourhood of zero

Theorem 1 *If* $\mathbb{E}[\xi_0] < \infty$ *&* $\mathbb{P}\{\xi_0 = 0\} < 1$, *the trajectories of* ρ *are* $a.s.$ *well-defined and non-defective hazard rates:*

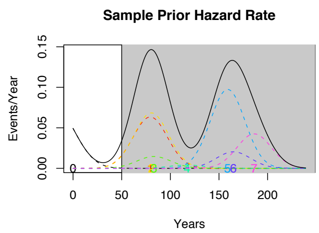
$$\exists t > 0 : \int_0^t \rho(s)ds < \infty \quad \& \quad \int_0^{\infty} \rho(s)ds = \infty.$$

Remark. In particular, this shows that the construction is **valid** if ξ_0 follows a *gamma* distribution (conjugate choice)

Theorem 2 *Let both* k_0 *and* k *be* r *times continuously differentiable on their domains. Furthermore, let* $k^{(i)}$*, the i-th derivative of* k*, be integrable on* $\mathbb{R}$ *and such that* $k^{(i)}(x) \downarrow 0$*, as* $x \to -\infty$*. Then,* $a.s.$ *the trajectories of* ρ *are* r *times continuously differentiable on* $\mathbb{R}_+$.

Remark. For example, if k is a zero mean normal probability density, the construction gives infinitely **smooth** hazard rates

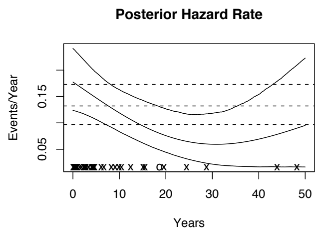
The proposed hazard rate construction can be interpreted in terms of countably many (defective) **competing hazard sources**; this gives insight into the prior distribution...



...and leads to a **straightforward MCMC approximation** of the posterior distribution.

A time-scale equivariant procedure is given to express **weak prior opinions** as follows:

- a prior pointwise expected hazard rate is imposed by suitably choosing k_0, so that
$$\mathbb{E}[\rho(t)] \equiv r_0$$
where r_0 is given by prior knowledge
- prior variability is controlled by letting
$$\sqrt{\lim_{t\to\infty} \mathbb{V}\text{ar}[\rho(t)]} = Hr_0$$
where H should be "big enough"
- prior oscillations are controlled by letting
$$T_\infty \sqrt{\lim_{t\to\infty} \mathbb{E}[\rho'(t)^2]} = 2(Hr_0)M_\infty$$
where T_∞ is a time-horizon of interest and M_∞ is a prior guess of the number of *extremes* in $[0, T_\infty]$



Pointwise expected value together with 2.5% and 97.5% quantiles (95% credible interval)

Solid lines refer to proposed prior, dashed lines to non-informative conjugate gamma prior for exponential inter-event times

The first 46 inter-event times (exact) are marked with X, the last one (right censored) is marked with O

References

- Dykstra R.L. and Laud P. (1981)
The Annals of Statistics, 9, 356–367.
- Padgett W.J. and Wei L.J. (1981)
Communications in Statistics, Part A, Theory and Methods, 10, 49–63.
- Lo A.Y. and Weng C.S. (1989)
Annals of the Institute of Statistical Mathematics, 41, 227–245.
- Ishwaran H. and James L.F. (2004)
Journal of the American Statistical Association, 99, 175–190.