Multiple change-point analysis

A Bayesian approach to change-point analysis for point processes $(y_1, y_2, \ldots, y_n)$: combines

- Poisson-process likelihood: $p(y|x) \propto \exp\{\sum_{i=1}^{n} \log x(y_i) - \int_0^L x(t)dt\}$
- prior model for step function $x(t)$, $0 \le t < L$, representing intensity

Prior model: represent step function by $(k, \{s_j\}_{j=1}^k, \{h_j\}_{j=0}^k)$: $x(t) = \sum_j h_j I_{[s_j, s_{j+1})}(t)$:

- number of steps k: Poisson(λ),
- step heights h_j: Gamma(α, β),
- step positions s_j: $p(s|k) \propto \prod_j (s_{j+1} - s_j)$,

all independent.

MCMC for step functions

$$\boldsymbol{\theta} = (k, \{s_j\}_{j=1}^k, \{h_j\}_{j=0}^k, \alpha, \beta)$$

I will use four moves:

(a) Metropolis change to a randomly chosen step height h_j.

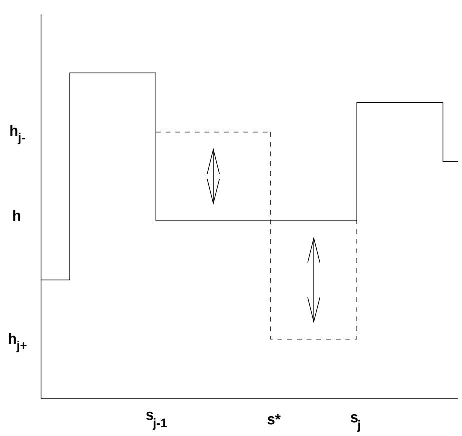
(b) Metropolis change to a randomly chosen step position s_j.

(c) Jump move: birth / death of steps

- birth: choose new step position s^* at random, split current step height h into two: (h_-, h_+)
- death: choose step at random to kill, combine current step heights (h_-, h_+) into one: h

(d) Update hyperparameters α, β

Birth and death of steps

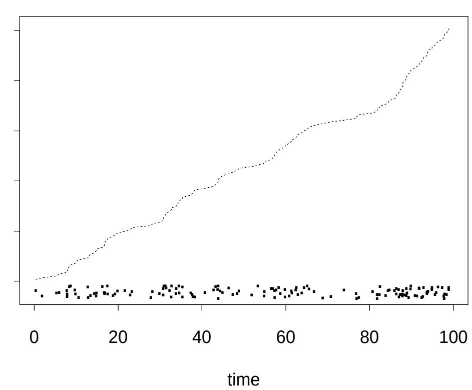


$$h_-^{w_-} h_+^{w_+} = h^{w_- + w_+}$$

$$(h, w, s^\star, u) \leftrightarrow (h_-, h_+, w_-, w_+)$$

Example: cyclones hitting the Bay of Bengal

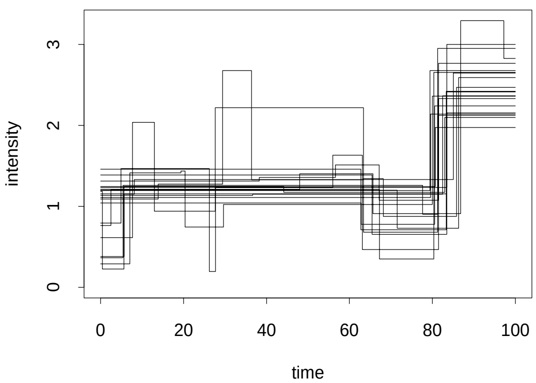
141 cyclones over a period of 100 years
(a cyclone is a storm with winds > 88 km h^{-1}).



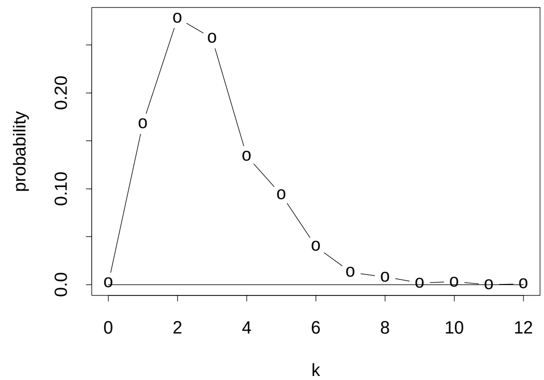
Choices of hyperparameters:

- Prior on k: Poisson(λ), with $\lambda = 3$.
- Prior on h_j: Gamma(α,β), with
  - $\alpha \sim \Gamma(2, 2)$
  - $\beta \sim \Gamma(1, n/L)$

Sample of step functions from the posterior:

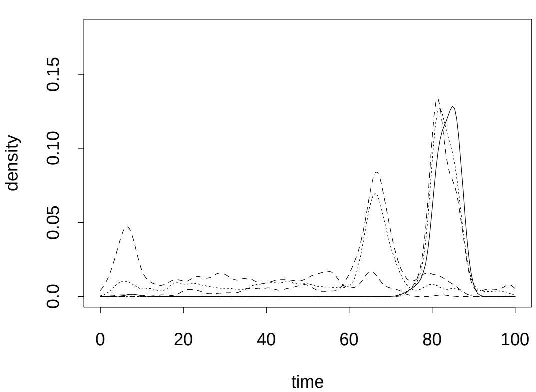


Posterior for the number of change points k

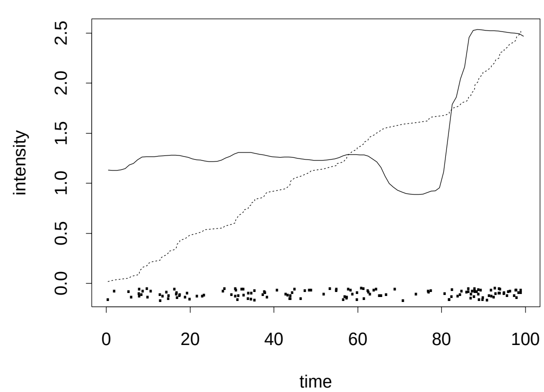


Zero change points is ruled out; $k = 1$ or 2 more probable than under the prior.

Posterior density estimates for change-point positions

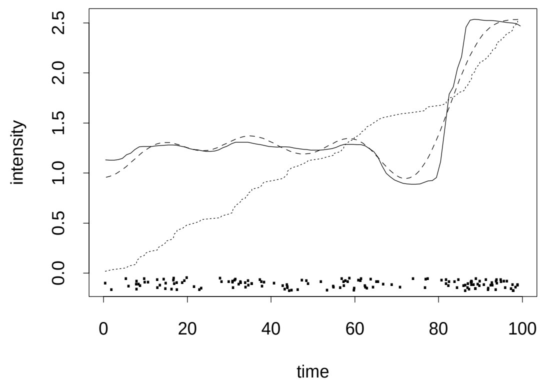


Model-averaged estimate: $E(x(\cdot)|y)$



(the expectation of a random step function is not a step function).

Ordinary smoothing methods (in this case a kernel smoother) can't match that mean curve



– fixed-bandwidth smoothers either over-smooth the steps, or under-smooth the plateaux.