Change-point detection in a Poisson process

S. Robin

joint work with E. Lebarbier, C. Dion-Blanc [DBLR23]

Sorbonne université

Math pour les Sciences du Vivant, Jan. 2024

Examples

Bat cries (night of the 17 jul. 2019)

Examples

Point process on $t \in [0, 1]$.

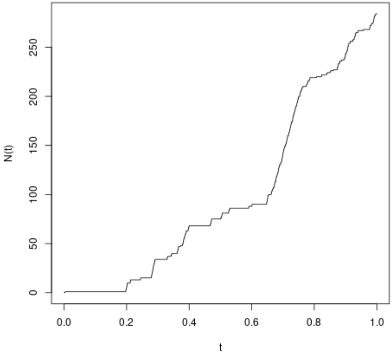
Event times:

$$0 < T_1 < \dots T_i < \dots T_n < 1$$

Counting process:

$$N(t) = \sum_{i=1}^{n} \mathbb{I}\{T_i \leqslant t\}$$

Bat cries (night of the 17 jul. 2019)[a]



[a]source: Vigie-Chiro program, Y. Bas, CESCO-MNHN

Examples

Point process on $t \in [0, 1]$.

Event times:

$$0 < T_1 < \dots T_i < \dots T_n < 1$$

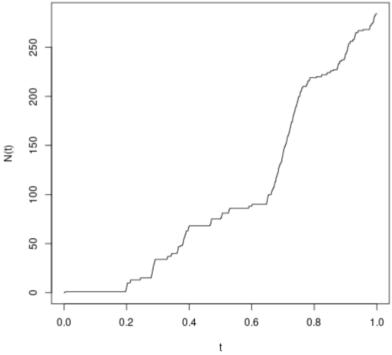
Counting process:

$$N(t) = \sum_{i=1}^{n} \mathbb{I}\{T_i \leqslant t\}$$

Poisson Process.

$$\{N(t)\}_{0 \leqslant t \leqslant 1} \sim PP(\lambda(t))$$

Bat cries (night of the 17 jul. 2019)[a]



[a] source: Vigie-Chiro program, Y. Bas, CESCO-MNHN

Intensity function $\lambda(t)$:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\mathbb{P}\{N(t + \Delta t) - N(t) = 1\}}{\Delta t}, \qquad \mathbb{E}N(s) - \mathbb{E}N(t) = \int_t^s \lambda(u)\,\mathrm{d}u$$

Change-point detection

Piecewise constant intensity function.

Change-points

$$(\tau_0 =)\ 0 < \tau_1 \cdots < \tau_{K-1} < 1\ (= \tau_K)$$

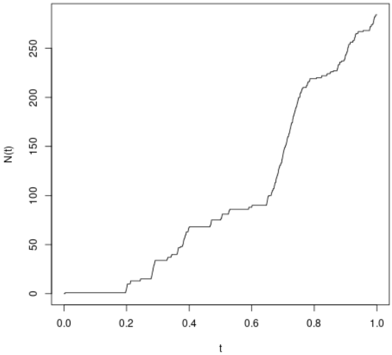
For $t \in I_k =]\tau_{k-1}, \tau_k]$:

$$\lambda(t) = \lambda_k$$

$\rightarrow$ Continuous piecewise linear cumulated intensity function

$$\Lambda(0, t) = \int_0^t \lambda(s)\ ds.$$

Bat cries (night of the 17 jul. 2019)[a]



[a] source: Vigie-Chiro program, Y. Bas, CESCO-MNHN

Change-point detection

Piecewise constant intensity function.

Change-points

$$(\tau_0 =)\ 0 < \tau_1 \cdots < \tau_{K-1} < 1\ (= \tau_K)$$

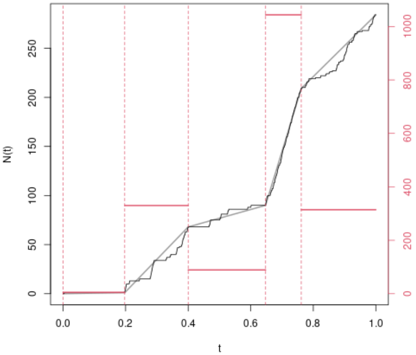
For $t \in I_k =]\tau_{k-1}, \tau_k]$:

$$\lambda(t) = \lambda_k$$

$\rightarrow$ Continuous piecewise linear cumulated intensity function

$$\Lambda(0, t) = \int_0^t \lambda(s)\ \mathrm{d}s.$$

Bat cries (night of the 17 jul. 2019)[a]



[a] source: Vigie-Chiro program, Y. Bas, CESCO-MNHN

Aim.

- Segmentation: estimate (τ, λ) in a reasonably fast manner
- Model selection: choose K

Change-point detection

Piecewise constant intensity function.

Change-points

$$(\tau_0 =)\ 0 < \tau_1 \cdots < \tau_{K-1} < 1\ (= \tau_K)$$

For $t \in I_k =]\tau_{k-1}, \tau_k]$:

$$\lambda(t) = \lambda_k$$

$\rightarrow$ Continuous piecewise linear cumulated intensity function

$$\Lambda(0, t) = \int_0^t \lambda(s)\ \mathrm{d}s.$$

Kilauea eruptions

Aim.

- Segmentation: estimate (τ, λ) in a reasonnably fast manner
- Model selection: choose K

Change-point detection

Piecewise constant intensity function.

Change-points

$$(\tau_0 =)\ 0 < \tau_1 \cdots < \tau_{K-1} < 1\ (= \tau_K)$$

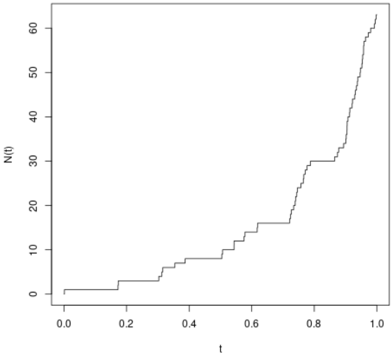
For $t \in I_k =]\tau_{k-1}, \tau_k]$:

$$\lambda(t) = \lambda_k$$

$\rightarrow$ Continuous piecewise linear cumulated intensity function

$$\Lambda(0, t) = \int_0^t \lambda(s)\ \mathrm{d}s.$$

Kilauea eruptions (from 1750 to 1984)[a]



[a] source: [HB17]

Aim.

- Segmentation: estimate (τ, λ) in a reasonably fast manner
- Model selection: choose K

Change-point detection

Piecewise constant intensity function.

Change-points

$$(\tau_0 =)\ 0 < \tau_1 \cdots < \tau_{K-1} < 1\ (= \tau_K)$$

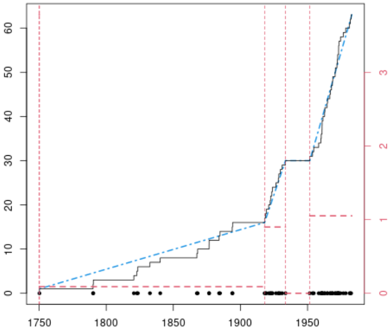
For $t \in I_k =]\tau_{k-1}, \tau_k]$:

$$\lambda(t) = \lambda_k$$

$\rightarrow$ Continuous piecewise linear cumulated intensity function

$$\Lambda(0, t) = \int_0^t \lambda(s)\ \mathrm{d}s.$$

Kilauea eruptions (from 1750 to 1984)[a]



[a] source: [HB17]

Aim.

- Segmentation: estimate (τ, λ) in a reasonably fast manner
- Model selection: choose K

Change-point detection

Three typical steps.

1. Propose a set of reasonably realistic models;
2. Design an (efficient) algorithm to get the parameter estimates;
3. Choose among the models.

Change-point detection

Three typical steps.

1. Propose a set of reasonably realistic models;
2. Design an (efficient) algorithm to get the parameter estimates;
3. Choose among the models.

Example.

1. $N(t)$ is a Poisson process with piece-wise constant intensity function λ:
$$\lambda(t) = \lambda_k \qquad \text{if} \quad \tau_{k-1} \leqslant t < \tau_k.$$
Parameters: change-points $\tau = (\tau_k)_{1 \leqslant k \leqslant K-1}$ and intensities $\lambda = (\lambda_k)_{1 \leqslant k \leqslant K}$.
2. For a given number of segments K find
$$(\widehat{\lambda}, \widehat{\tau}) = \arg\min_{\tau, \lambda} C_K(N; \tau, \lambda), \qquad \text{e.g.} \quad C_K(N; \tau, \lambda) = -\log p_{K,\tau,\lambda}(N);$$
3. Choose the number of segments K.

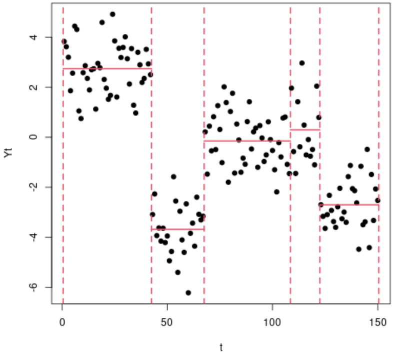
Outline

Discrete time

A simple discrete-time problem.

- Data: $Y = (Y_t)_{t=1,\dots n}$ independent;
- Change-points: $\tau = (\tau_1, \dots \tau_{K-1})$;
- Means: $\mu = \mu_1, \dots \mu_K$;

$$\tau_{k-1} < t \leqslant \tau_k \quad \Rightarrow \quad Y_t \sim \mathcal{N}(\mu_k, 1).$$

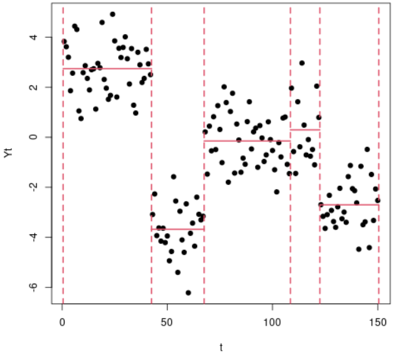


Discrete time

A simple discrete-time problem.

- Data: $Y = (Y_t)_{t=1,\dots n}$ independent;
- Change-points: $\tau = (\tau_1, \dots \tau_{K-1})$;
- Means: $\mu = \mu_1, \dots \mu_K$;

$$\tau_{k-1} < t \leqslant \tau_k \quad \Rightarrow \quad Y_t \sim \mathcal{N}(\mu_k, 1).$$



Aim. Estimate τ and μ from Y.

Specificity. μ is continuous whereas τ is discrete: $\mu \in \mathbb{R}^K$, $\tau \in [\![n-1]\!]^{K-1}$.

Maximum-likelihood inference

Principle. For a given number of segments K, look for

$$\widehat{\theta} = \arg\max_{\theta} \log p_{\theta}(Y) = \arg\min_{\theta} \sum_{k=1}^{K} \sum_{t=\tau_{k-1}+1}^{\tau_k} (Y_t - \mu_k)^2$$

where $\theta = (\tau, \mu)$.

Maximum-likelihood inference

Principle. For a given number of segments K, look for

$$\widehat{\theta} = \arg\max_{\theta} \log p_\theta(Y) = \arg\min_{\theta} \sum_{k=1}^{K} \sum_{t=\tau_{k-1}+1}^{\tau_k} (Y_t - \mu_k)^2$$

where $\theta = (\tau, \mu)$.

Estimating μ. For a given τ, we have that

$$\widehat{\mu}_k(\tau) = \arg\min_{\mu_k} \sum_{t=\tau_{k-1}+1}^{\tau_k} (Y_t - \mu_k)^2 = \frac{1}{\tau_k - \tau_{k-1}} \sum_{t=\tau_{k-1}+1}^{\tau_k} Y_t.$$

Maximum-likelihood inference

Principle. For a given number of segments K, look for

$$\widehat{\theta} = \arg\max_{\theta} \log p_\theta(Y) = \arg\min_{\theta} \sum_{k=1}^{K} \sum_{t=\tau_{k-1}+1}^{\tau_k} (Y_t - \mu_k)^2$$

where $\theta = (\tau, \mu)$.

Estimating μ. For a given τ, we have that

$$\widehat{\mu}_k(\tau) = \arg\min_{\mu_k} \sum_{t=\tau_{k-1}+1}^{\tau_k} (Y_t - \mu_k)^2 = \frac{1}{\tau_k - \tau_{k-1}} \sum_{t=\tau_{k-1}+1}^{\tau_k} Y_t.$$

Estimating τ. We are left with the discrete optimization problem

$$\widehat{\tau} = \arg\max_{\tau} \log p_{(\tau, \widehat{\mu}(\tau))}(Y) = \arg\min_{\tau} \sum_{k=1}^{K} \underbrace{\sum_{t=\tau_{k-1}+1}^{\tau_k} (Y_t - \widehat{\mu}_k(\tau))^2}_{C(\tau_{k-1}+1, \tau_k)}$$

Discrete-time segmentation problem

- Consider a cost function $C(t_1, t_2)$ defined for $1 \leqslant t_1 < t_2 \leqslant n$;
- Define the segmentation space

$$\mathcal{T}_K = \{\tau \in [\![n-1]\!]^{K-1} : 1 \leqslant \tau_1 < \dots \tau_{K-1} < n\},$$

  observe that

$$\text{card}(\mathcal{T}) = \binom{n-1}{K-1} = \mathcal{O}(n^K);$$

- Define the contrast $\gamma : \mathcal{T} \mapsto \mathbb{R}$:

$$\gamma(\tau) = \sum_{k=1}^{K} C(\tau_{k-1}+1, \tau_k), \qquad \text{with } \tau_0 = 0, \tau_K = n.$$

Discrete-time segmentation problem

- Consider a cost function $C(t_1, t_2)$ defined for $1 \leqslant t_1 < t_2 \leqslant n$;
- Define the segmentation space

$$\mathcal{T}_K = \{\tau \in [\![n-1]\!]^{K-1} : 1 \leqslant \tau_1 < \dots \tau_{K-1} < n\},$$

  observe that

$$\mathrm{card}(\mathcal{T}) = \binom{n-1}{K-1} = \mathcal{O}(n^K);$$

- Define the contrast $\gamma : \mathcal{T} \mapsto \mathbb{R}$:

$$\gamma(\tau) = \sum_{k=1}^{K} C(\tau_{k-1}+1, \tau_k), \qquad \text{with } \tau_0 = 0, \tau_K = n.$$

Dynamic programming. The optimal segmentation

$$\widehat{\tau} = \arg\min_{\tau \in \mathcal{T}_K} \gamma(\tau)$$

can be recovered in $\mathcal{O}(n^2)$ using a dynamic programming algorithm.

Dynamic programming

Belman's principle. All sub-paths of the optimal path must be optimal themselves.

Dynamic programming

Belman's principle. All sub-paths of the optimal path must be optimal themselves.

Dynamic programming algorithm [AL89].

1. For each $m \in [\![n]\!]$, compute (in $\mathcal{O}(m)$)

$$t_1(m) = \underset{1 \leqslant t < m}{\arg\min}\, C(1, t) + C(t+1, m), \qquad S_2(m) = \min_{1 \leqslant t < m}();$$

2. Then, for each $k = 3 \dots K$ and each $m \in [\![n]\!]$, compute (in $\mathcal{O}(m)$)

$$t_{k-1}(m) = \underset{k-1 \leqslant t < m}{\arg\min}\, S_{k-1}(1, t) + C(t+1, m), \qquad S_k(m) = \min_{k-1 \leqslant t < m}().$$

Dynamic programming

Belman's principle. All sub-paths of the optimal path must be optimal themselves.

Dynamic programming algorithm [AL89].

1. For each $m \in [\![n]\!]$, compute (in $\mathcal{O}(m)$)

$$t_1(m) = \arg\min_{1 \leqslant t < m} C(1,t) + C(t+1,m), \qquad S_2(m) = \min_{1 \leqslant t < m} ();$$

2. Then, for each $k = 3 \dots K$ and each $m \in [\![n]\!]$, compute (in $\mathcal{O}(m)$)

$$t_{k-1}(m) = \arg\min_{k-1 \leqslant t < m} S_{k-1}(1,t) + C(t+1,m), \qquad S_k(m) = \min_{k-1 \leqslant t < m} ().$$

Solution. After $\mathcal{O}(Kn^2)$ operations:

$$\min_{\tau \in \mathcal{T}_K} \gamma(\tau) = S_K(n),$$

and

$$\widehat{\tau}_{K-1} = t_{K-1}(n), \qquad \widehat{\tau}_{k-1} = t_{K-1}(\widehat{\tau}_k).$$

Outline

Maximum-likelihood segmentation

First useful property of Poisson processes: Independence of disjoint intervals.

Maximum-likelihood segmentation

First useful property of Poisson processes: Independence of disjoint intervals.

(Neg-log-)likelihood. Denoting

- $\Delta\tau_k$ the length of the k-th interval $(= \tau_k - \tau_{k-1})$,
- ΔN_k the number of events within the k-th interval $(= N(\tau_k) - N(\tau_{k-1}))$:

$$-\log p_{\tau,\lambda}(N) = \sum_{k=1}^{K} \lambda_k \Delta\tau_k - \Delta N_k \log \lambda_k,$$

Maximum-likelihood segmentation

First useful property of Poisson processes: Independence of disjoint intervals.

(Neg-log-)likelihood. Denoting

- $\Delta\tau_k$ the length of the k-th interval ($= \tau_k - \tau_{k-1}$),
- ΔN_k the number of events within the k-th interval ($= N(\tau_k) - N(\tau_{k-1})$):

$$-\log p_{\tau,\lambda}(N) = \sum_{k=1}^{K} \lambda_k \Delta\tau_k - \Delta N_k \log \lambda_k,$$

Additive contrast. General form = sum over the segments

$$\gamma(\tau, \lambda) = \sum_{k=1}^{K} C(\Delta N_k, \Delta\tau_k, \lambda_k)$$

Maximum-likelihood segmentation

First useful property of Poisson processes: Independence of disjoint intervals.

(Neg-log-)likelihood. Denoting

- $\Delta\tau_k$ the length of the k-th interval $(= \tau_k - \tau_{k-1})$,
- ΔN_k the number of events within the k-th interval $(= N(\tau_k) - N(\tau_{k-1}))$:

$$-\log p_{\tau,\lambda}(N) = \sum_{k=1}^{K} \lambda_k \Delta\tau_k - \Delta N_k \log \lambda_k,$$

Additive contrast. General form = sum over the segments

$$\gamma(\tau, \lambda) = \sum_{k=1}^{K} C(\Delta N_k, \Delta\tau_k, \lambda_k)$$

Optimization problem.

$$(\widehat{\tau}, \widehat{\lambda}) = \underset{\tau \in \mathcal{T}_K, \lambda \in (\mathbb{R}^+)^K}{\arg\min} \quad \gamma(\tau, \lambda).$$

Minimizing the contrast function

Optimal λ. Because the contrast is additive, we may define

$$\widehat{\lambda}_k = \widehat{\lambda}_k(\tau) = \arg\min_{\lambda_k \in \mathbb{R}^+} C(\Delta N_k, \Delta \tau_k, \lambda_k)$$

e.g. $\widehat{\lambda}_k = \Delta N_k / \Delta \tau_k$ if $\gamma = -\log p_\theta$.

Minimizing the contrast function

Optimal λ. Because the contrast is additive, we may define

$$\widehat{\lambda}_k = \widehat{\lambda}_k(\tau) = \arg\min_{\lambda_k \in \mathbb{R}^+} C(\Delta N_k, \Delta\tau_k, \lambda_k)$$

e.g. $\widehat{\lambda}_k = \Delta N_k / \Delta\tau_k$ if $\gamma = -\log p_\theta$.

Optimal τ. We are left with the minimization problem

$$\widehat{\tau} = \arg\min_{\tau \in \mathcal{T}_K} \; \widehat{\gamma}(\tau), \qquad \text{where} \quad \widehat{\gamma}(\tau) = \gamma(\tau, \widehat{\lambda}(\tau))$$

where $\mathcal{T}_K$ is the continuous segmentation space:

$$\mathcal{T}_K = \left\{\tau \in [0,1]^{K+1} : 0 = \tau_0 < \tau_1 \cdots < \tau_{K-1} < \tau_K = 1\right\}.$$

Minimizing the contrast function

Optimal λ. Because the contrast is additive, we may define

$$\widehat{\lambda}_k = \widehat{\lambda}_k(\tau) = \arg\min_{\lambda_k \in \mathbb{R}^+} C(\Delta N_k, \Delta\tau_k, \lambda_k)$$

e.g. $\widehat{\lambda}_k = \Delta N_k / \Delta\tau_k$ if $\gamma = -\log p_\theta$.

Optimal τ. We are left with the minimization problem

$$\widehat{\tau} = \arg\min_{\tau \in \mathcal{T}_K} \widehat{\gamma}(\tau), \qquad \text{where} \quad \widehat{\gamma}(\tau) = \gamma(\tau, \widehat{\lambda}(\tau))$$

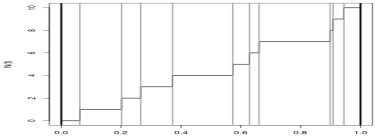
where $\mathcal{T}_K$ is the continuous segmentation space:

$$\mathcal{T}_K = \left\{ \tau \in [0,1]^{K+1} : 0 = \tau_0 < \tau_1 \cdots < \tau_{K-1} < \tau_K = 1 \right\}.$$

Main issue: The contrast $\widehat{\gamma}(\tau)$ is neither convex nor continuous wrt τ.

Shape of the contrast fonction

Observed $N(t)$: $n = 10$,



[1] gray borders come by pair

Shape of the contrast fonction

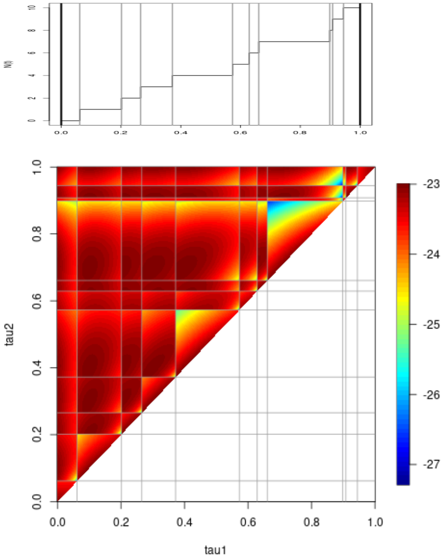
Observed $N(t)$: $n = 10$,

Contrast $\widehat{\gamma}(\tau)$ for $K = 3$ segments:

$$\tau = (\tau_1, \tau_2).$$

One 'block' =
one specific value for the vector [1]

$$\Delta N = (\Delta N_1, \Delta N_2, \Delta N_3)$$



[1] gray borders come by pair

Partitioning the segmentation space

Partitioning the number of events. Define $\mathcal{N}^K = \left\{\nu \in \mathbb{N}^K : \sum_{k=1}^K \nu_k = n\right\}$.

$\rightarrow$ ν_k = given number of events in segment k.

Partitioning the segmentation space

Partitioning the number of events. Define $\mathcal{N}^K = \left\{\nu \in \mathbb{N}^K : \sum_{k=1}^K \nu_k = n\right\}$.
$\rightarrow$ ν_k = given number of events in segment k.

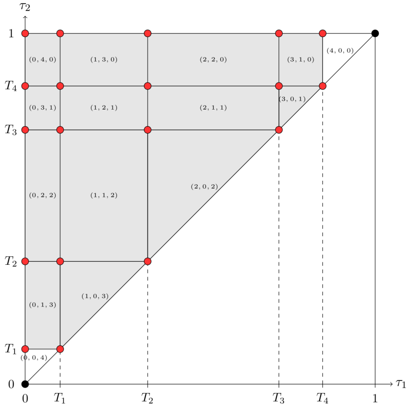
Partitioning the segmentation space. For $\nu \in \mathcal{N}_K$, define

$$\mathcal{T}(\nu) = \{\tau \in \mathcal{T}_K : \Delta N = \nu\}.$$

$\rightarrow$ $\mathcal{T}(\nu)$ = set of segmentation satisfying the prescribed $\nu = (\nu_1, \dots \nu_K)$.

We have

$$\min_{\tau \in \mathcal{T}_K} \widehat{\gamma}(\tau) = \min_{\nu \in \mathcal{N}^K} \min_{\tau \in \mathcal{T}(\nu)} \widehat{\gamma}(\tau,).$$



Optimal segmentation

Proposition 1. If $K \leqslant n$ and if $\widehat{\gamma}(\tau)$ is strictly concave wrt $\tau \in \mathcal{T}(\nu)$ for each $\nu \in \mathcal{N}^K$, then

$$\widehat{\tau} = \arg\min_{\tau \in \mathcal{T}_K} \widehat{\gamma}(\tau) \subset \{T_1^-, T_1, T_2^-, T_2^-, \dots T_n^-, T_n\}.$$

Optimal segmentation

Proposition 1. If $K \leqslant n$ and if $\widehat{\gamma}(\tau)$ is strictly concave wrt $\tau \in \mathcal{T}(\nu)$ for each $\nu \in \mathcal{N}^K$, then

$$\widehat{\tau} = \arg\min_{\tau \in \mathcal{T}_K} \widehat{\gamma}(\tau) \subset \{T_1^-, T_1, T_2^-, T_2^-, \dots T_n^-, T_n\}.$$

Proposition 2. If each $\widehat{C}(\nu_k, \Delta\tau_k) := C(\nu_k, \Delta\tau_k, \widehat{\lambda}_k)$ is strictly concave wrt $\Delta\tau_k$, then $\widehat{\gamma}(\tau)$ is strictly concave wrt $\tau \in \mathcal{T}(\nu)$.

Optimal segmentation

Proposition 1. If $K \leqslant n$ and if $\widehat{\gamma}(\tau)$ is strictly concave wrt $\tau \in \mathcal{T}(\nu)$ for each $\nu \in \mathcal{N}^K$, then

$$\widehat{\tau} = \arg\min_{\tau \in \mathcal{T}_K} \widehat{\gamma}(\tau) \subset \{T_1^-, T_1, T_2^-, T_2^-, \dots T_n^-, T_n\}.$$

Proposition 2. If each $\widehat{C}(\nu_k, \Delta\tau_k) := C(\nu_k, \Delta\tau_k, \widehat{\lambda}_k)$ is strictly concave wrt $\Delta\tau_k$, then $\widehat{\gamma}(\tau)$ is strictly concave wrt $\tau \in \mathcal{T}(\nu)$.

Consequence. $\widehat{\tau}$ can be obtained by dynamic programming over the $2n + 2$ possible change-points

$$\mathcal{S} = \{0, T_1^-, T_1, T_2^-, T_2, \dots T_n^-, T_n, 1\}$$

with complexity at most $O(n^2)$.

Admissible contrasts

Poisson contrast. $\widehat{C}_P(\nu_k, \Delta\tau_k) = \nu_k(1 - \log \nu_k + \log \Delta\tau_k)$ is concave wrt $\Delta\tau$.

Admissible contrasts

Poisson contrast. $\widehat{C}_P(\nu_k, \Delta\tau_k) = \nu_k(1 - \log \nu_k + \log \Delta\tau_k)$ is concave wrt $\Delta\tau$.

Poisson-Gamma model. For each segment $1 \leqslant k \leqslant K$:

$$\Lambda_k \text{ iid } \sim \mathcal{G}\text{am}(a, b), \qquad \{N(t)\}_{t \in I_k} \mid \Lambda_k \sim PP(\Lambda_k).$$

Admissible contrasts

Poisson contrast. $\widehat{C}_P(\nu_k, \Delta\tau_k) = \nu_k(1 - \log \nu_k + \log \Delta\tau_k)$ is concave wrt $\Delta\tau$.

Poisson-Gamma model. For each segment $1 \leqslant k \leqslant K$:

$$\Lambda_k \text{ iid } \sim \mathcal{G}\text{am}(a, b), \qquad \{N(t)\}_{t \in I_k} \mid \Lambda_k \sim PP(\Lambda_k).$$

Contrast for one segment:

$$C_{PG}(\Delta N_k, \Delta\tau_k) = \text{cst} - \log \Gamma(a + \Delta N_k) + (a + \Delta N_k) \log(b + \Delta\tau_k)$$

$\rightarrow$ Strictly concave wrt $\Delta\tau_k$.

Desirable contrast

Remark. The Poisson contrast $\widehat{C}_P(\nu_k, \Delta\tau_k) = \nu_k(1 - \log \nu_k + \log \Delta\tau_k)$ satisfies

$$\widehat{C}_P(\nu_k = 1, \Delta\tau_k = 0) = -\infty.$$

- The optimal solution will involve segments with null length and containing only one event.
- 'Undesirable' contrast.

Desirable contrast

Remark. The Poisson contrast $\widehat{C}_P(\nu_k, \Delta\tau_k) = \nu_k(1 - \log \nu_k + \log \Delta\tau_k)$ satisfies

$$\widehat{C}_P(\nu_k = 1, \Delta\tau_k = 0) = -\infty.$$

- The optimal solution will involve segments with null length and containing only one event.
- 'Undesirable' contrast.

Poisson-Gamma contrast. $C_{PG}(\nu_k, \Delta\tau_k) = -\log \Gamma(a + \nu_k) + (a + \nu_k) \log(b + \Delta\tau_k)$.

- Satisfies the concavity property ($\rightarrow$ admissible),
- but avoids segments with null length ($\rightarrow$ desirable).

Outline

Model selection

Second useful property of Poisson processes: Thining.

- $\{N(t)\} \sim PP(\lambda(t))$
- Sample event times (with prob. f)
- Store the remaining events

Model selection

Second useful property of Poisson processes: Thining.

- $\{N(t)\} \sim PP(\lambda(t))$

- Sample event times (with prob. f)
- Store the remaining events

$$\{N^L(t)\} \sim PP(f\lambda(t)), \qquad \{N^T(t)\} \sim PP((1-f)\lambda(t)), \qquad \{N^L(t)\} \perp \{N^T(t)\}$$

Model selection

Second useful property of Poisson processes: Thining.

- $\{N(t)\} \sim PP(\lambda(t))$
- Sample event times (with prob. f)
- Store the remaining events

$$\{N^L(t)\} \sim PP(f\lambda(t)), \qquad \{N^T(t)\} \sim PP((1-f)\lambda(t)), \qquad \{N^L(t)\} \perp \{N^T(t)\}$$

Consequence. If $\{N(t)\}_{0 \leqslant t \leqslant 1} \sim PP(\lambda(t))$, with $\lambda(t)$ piecewise constant with change-points $\tau = (\tau_k)$ and intensities $\lambda = (\lambda_k)$, then

Model selection

Second useful property of Poisson processes: Thining.

- $\{N(t)\} \sim PP(\lambda(t))$
- Sample event times (with prob. f)
- Store the remaining events

$$\{N^L(t)\} \sim PP(f\lambda(t)), \qquad \{N^T(t)\} \sim PP((1-f)\lambda(t)), \qquad \{N^L(t)\} \perp \{N^T(t)\}$$

Consequence. If $\{N(t)\}_{0 \leqslant t \leqslant 1} \sim PP(\lambda(t))$, with $\lambda(t)$ piecewise constant with change-points $\tau = (\tau_k)$ and intensities $\lambda = (\lambda_k)$, then

- $\lambda^L(t)$ piecewise constant with change-points (τ_k) and intensities $(f\lambda_k)$,
- $\lambda^T(t)$ piecewise constant with change-points (τ_k) and intensities $((1-f)\lambda_k)$,
- $\{N^L(t)\} \perp \{N^T(t)\}$.

Cross validation

Sampling event times provides two independent Poisson processes with same change-points.

Cross validation

Sampling event times provides two independent Poisson processes with same change-points.

Cross-validation procedure. For $1 \leqslant K \leqslant K_{\max}$,

- Repeat for $1 \leqslant m \leqslant M$:

  1 – Sample the event times to form $\{N^{L,m}(t)\}$ (learn) and $\{N^{T,m}(t)\}$ (test),

  2 – Estimate $\widehat{\tau}^{L,m}$ and $\widehat{\lambda}^{L,m}$ from $\{N^{L,m}(t)\}$,

  3 – Compute the contrast $\gamma_K^{T,m} = \gamma\left(\{N^T(t)\}; \widehat{\tau}^{L,m}, \frac{1-f}{f}\widehat{\lambda}^{L,m}\right)$.

Cross validation

Sampling event times provides two independent Poisson processes with same change-points.

Cross-validation procedure. For $1 \leqslant K \leqslant K_{\max}$,

- Repeat for $1 \leqslant m \leqslant M$:

  1 – Sample the event times to form $\{N^{L,m}(t)\}$ (learn) and $\{N^{T,m}(t)\}$ (test),

  2 – Estimate $\widehat{\tau}^{L,m}$ and $\widehat{\lambda}^{L,m}$ from $\{N^{L,m}(t)\}$,

  3 – Compute the contrast $\gamma_K^{T,m} = \gamma\left(\{N^T(t)\}; \widehat{\tau}^{L,m}, \dfrac{1-f}{f}\widehat{\lambda}^{L,m}\right)$.

- Compute

$$\overline{\gamma}_K = \frac{1}{M}\sum_{m=1}^{M} \gamma_K^{T,m}$$

Cross validation

Sampling event times provides two independent Poisson processes with same change-points.

Cross-validation procedure. For $1 \leqslant K \leqslant K_{\max}$,

- Repeat for $1 \leqslant m \leqslant M$:

  1 – Sample the event times to form $\{N^{L,m}(t)\}$ (learn) and $\{N^{T,m}(t)\}$ (test),

  2 – Estimate $\widehat{\tau}^{L,m}$ and $\widehat{\lambda}^{L,m}$ from $\{N^{L,m}(t)\}$,

  3 – Compute the contrast $\gamma_K^{T,m} = \gamma\left(\{N^T(t)\}; \widehat{\tau}^{L,m}, \frac{1-f}{f}\widehat{\lambda}^{L,m}\right)$.

- Compute

$$\overline{\gamma}_K = \frac{1}{M}\sum_{m=1}^{M} \gamma_K^{T,m}$$

- Select

$$\widehat{K} = \arg\min_K \overline{\gamma}_K$$

Outline

Practical implementation.

Contrasts. During the CV process, we use

- a sampling rate of $f = 4/5$,
- the Poisson-Gamma contrast γ_{PG} for the learning step and
- the Poisson contrast γ_P for the test step.

Practical implementation.

Contrasts. During the CV process, we use

- a sampling rate of $f = 4/5$,
- the Poisson-Gamma contrast γ_{PG} for the learning step and
- the Poisson contrast γ_P for the test step.

Hyper-parameters. For an observed path $\{N(t)\}_{0 \leqslant t \leqslant 1}$ with $n = N(1)$ events, we use

$$a = 1, \qquad b = 1/n$$

to fit the observed total number of events.

Practical implementation.

Contrasts. During the CV process, we use

- a sampling rate of $f = 4/5$,
- the Poisson-Gamma contrast γ_{PG} for the learning step and
- the Poisson contrast γ_P for the test step.

Hyper-parameters. For an observed path $\{N(t)\}_{0 \leqslant t \leqslant 1}$ with $n = N(1)$ events, we use

$$a = 1, \qquad b = 1/n$$

to fit the observed total number of events.

R package `CptPointProcess` available on `github.com/Elebarbier/CptPointProcess`.

Some simulations

Simulation setting. $K = 6$ segments with varying length. Tuning parameters:

- $\overline{\lambda}$ average intensity ($\rightarrow$ total number of events),
- λ_R = height of the steps ($\rightarrow$ contrast between segments).

Estimation. Choose K via CV, then refit the parameters to the whole dataset.

Some simulations

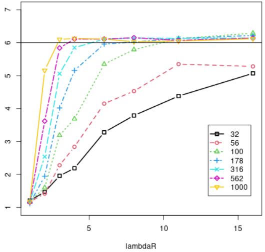
Simulation setting. $K = 6$ segments with varying length. Tuning parameters:

- $\overline{\lambda}$ average intensity ($\rightarrow$ total number of events),
- λ_R = height of the steps ($\rightarrow$ contrast between segments).

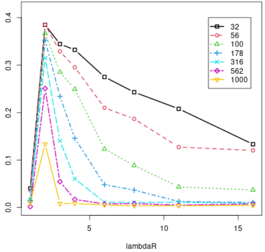
Estimation. Choose K via CV, then refit the parameters to the whole dataset.

Results.

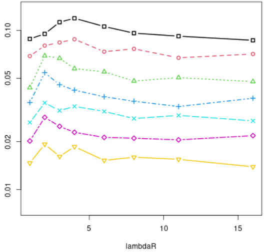
Model selection
$\widehat{K}$



Change-point location
Haussdorf$(\widehat{\tau}, \tau^*)$



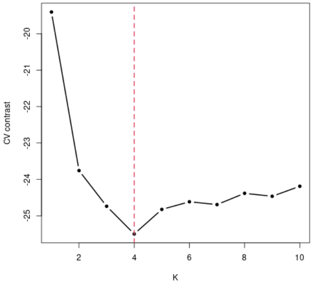
Cumulated intensity
$\ell_2(\widehat{\Lambda}, \Lambda^*)$



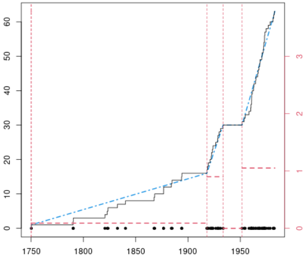
Kilauea eruptions

$n = 63$ eruptions reported between the mid 18th and the late 20th century.

Model selection via CV



Resulting segmentation



Outline

Reminder: segmentation in discrete time

Segmentation in continuous time

Model selection

Illustrations

Extension to marked-Poisson processes

Future works

Extension to marked-Poisson processes

Marked Poisson Process.

- $\{Y(t)\}_{0 \leqslant t \leqslant 1} \sim MPP(\lambda(t), \mu(t))$:

$$\{N(t)\}_{0 \leqslant t \leqslant 1} \sim PP(\lambda(t)), \qquad \text{at each } T_i: \quad X_i \sim \mathcal{F}(\mu(T_i))$$

- Volcanos: mark = eruption duration.
- Bat cries: mark = bat species or cry duration.

Extension to marked-Poisson processes

Marked Poisson Process.

- $\{Y(t)\}_{0 \leqslant t \leqslant 1} \sim MPP(\lambda(t), \mu(t))$:

$$\{N(t)\}_{0 \leqslant t \leqslant 1} \sim PP(\lambda(t)), \qquad \text{at each } T_i: \quad X_i \sim \mathcal{F}(\mu(T_i))$$

- Volcanos: mark = eruption duration.
- Bat cries: mark = bat species or cry duration.

Proposed method.

- Works the same way, provided that concavity holds.
- Poisson-Gamma events + Exponential-Gamma durations is both admissible and desirable.

Marked Poisson process: Etna eruptions

Count and marks: Events = eruptions, marks = duration of each eruption.

Model.

- Piecewise-constant intensity Poisson process for the events
- Exponential distribution (with segment specific parm.) for the durations

Marked Poisson process: Etna eruptions

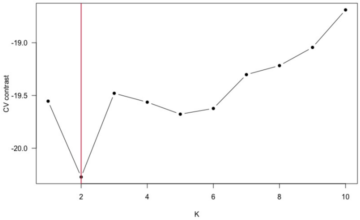
Count and marks: Events = eruptions, marks = duration of each eruption.

Model.

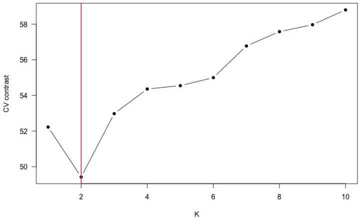
- Piecewise-constant intensity Poisson process for the events
- Exponential distribution (with segment specific parm.) for the durations

CV for the selection of K.

Poisson

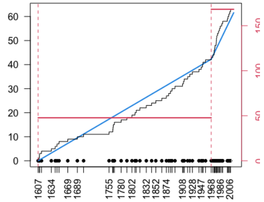


Marked Poisson



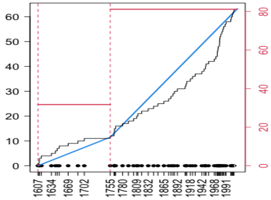
Marked Poisson process: Etna eruptions

Poisson ($\widehat{K} = 2$)
Events

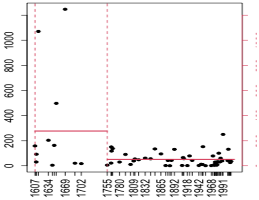


Marked Poisson ($\widehat{K} = 2$)

Events



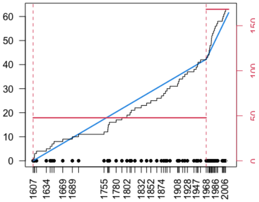
Marks



Marked Poisson process: Etna eruptions

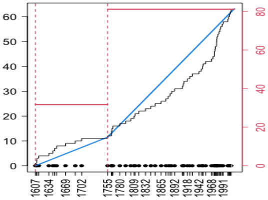
Poisson ($\widehat{K} = 2$)

Events

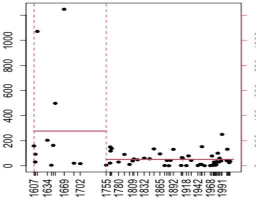


Marked Poisson ($\widehat{K} = 2$)

Events

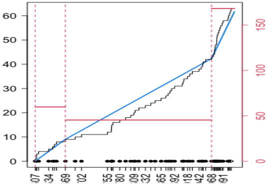


Marks

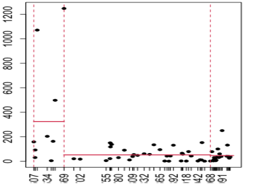


Marked Poisson ($K = 3$)

Events



Marks



Outline

Future works

Segmentation-clustering.

- Each segment belongs to a class $1 \leqslant q \leqslant Q$ (with probability π_q and intensity $\lambda_k = \ell_q$),
- Combination of EM and DP algorithms [PRLD07],
- Bat cries: Class = animal behaviour (hunt, transit, ...)

Future works

Segmentation-clustering.

- Each segment belongs to a class $1 \leqslant q \leqslant Q$ (with probability π_q and intensity $\lambda_k = \ell_q$),
- Combination of EM and DP algorithms [PRLD07],
- Bat cries: Class = animal behaviour (hunt, transit, ...)

And also.

- Theoretically grounded model selection criterion (BIC: ongoing),
- Consistency of the estimated change-points (ongoing),
- Other desirable contrasts, ...

References I

I. E. Auger and C. E. Lawrence. Algorithms for the optimal identification of segment neighborhoods. *Bull. Math. Biol.*, 51(1):39–54, 1989.

C. Dion-Blanc, E Lebarbier, and S Robin. Multiple change-point detection for Poisson processes. Technical Report 2302.09103, arXiv, 2023.

C-H Ho and M Bhaduri. A quantitative insight into the dependence dynamics of the Kilauea and Mauna Loa volcanoes, Hawaii. *Mathematical Geosciences*, 49(7):893–911, 2017.

F. Picard, S. Robin, E Lebarbier, and J-J Daudin. A segmentation/clustering model for the analysis of array CGH data. *Biometrics*, 63(3):758–766, 2007.

Appendix

Number of elements in the partition of the segmentation space.

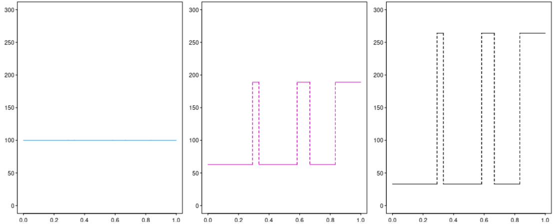
$$|\mathcal{N}_K| = \sum_{h=\lfloor (K-1)/2 \rfloor}^{K} \binom{n-1}{h-1} \binom{h+1}{K-h}$$

Appendix

Number of elements in the partition of the segmentation space.

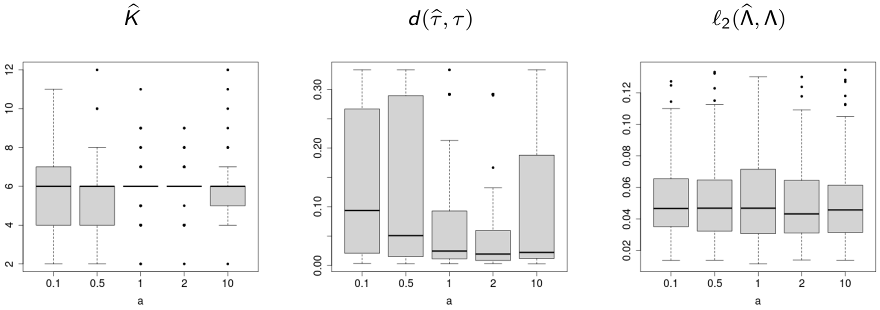
$$|\mathcal{N}_K| = \sum_{h=\lfloor (K-1)/2 \rfloor}^{K} \binom{n-1}{h-1} \binom{h+1}{K-h}$$

Simulations: Shape of the intensity function $\lambda(t)$. $K = 6$, $\overline{\lambda} = 100$, $\lambda_R = 1, 3, 8$.



Choice of the hyper-parameters (a, b)

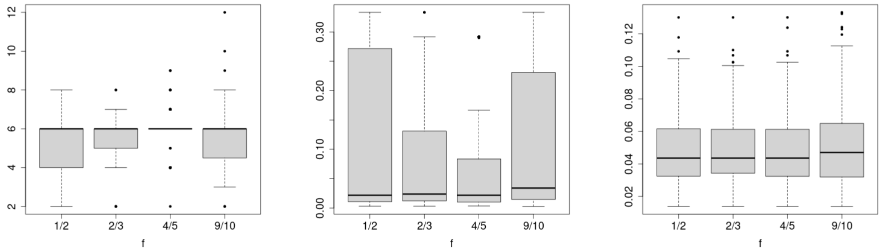
Sticking to $a/b = n$ ($\overline{\lambda} = 100$ and $\lambda_R = 8$) :



Choice of the fraction f for cross-validation

Proposed $f = 4/5$ ($\overline{\lambda} = 100$ and $\lambda_R = 8$) :

$\widehat{K}$ $\quad$ $d(\widehat{\tau}, \tau)$ $\quad$ $\ell_2(\widehat{\Lambda}, \Lambda)$



Poisson process: Mauna Loa eruptions