POLITECNICO
MILANO 1863

Bayesian Inference on Poisson-Gamma RMs for Spatial Statistics

Bachelor's Thesis in Mathematical Engineering

Author: Edoardo Tarcisio Visconti
Advisor: Prof. Federico Bassetti

18/07/2024

Biometrika (1998), **85**, 2, pp. 251–267
Printed in Great Britain

Poisson/gamma random field models for spatial statistics

BY ROBERT L. WOLPERT
Institute of Statistics and Decision Sciences, Duke University, Durham, North Carolina 27708-0251, U.S.A.
wolpert@stat.duke.edu

AND KATJA ICKSTADT
Department of Statistics, University of North Carolina, Chapel Hill, North Carolina 27599-3260, U.S.A.
katja@stat.unc.edu

SUMMARY

Doubly stochastic Bayesian hierarchical models are introduced to account for uncertainty and spatial variation in the underlying intensity measure for point process models. Inhomogeneous gamma process random fields and, more generally, Markov random fields with infinitely divisible distributions are used to construct positively autocorrelated intensity measures for spatial Poisson point processes; these in turn are used to model the number and location of individual events. A data augmentation scheme and Markov chain Monte Carlo numerical methods are employed to generate samples from Bayesian posterior and predictive distributions. The methods are developed in both continuous and discrete settings, and are applied to a problem in forest ecology.

Some key words: Bayesian mixture model; Bioabundance; Cox process; Data augmentation; Lévy process; Markov chain Monte Carlo; Simulation.

Abstract

1) Alternative approach based on discretizing the underlying Gamma process

2) Advantages:

- Simpler model
- Computationally less expensive
- More stable
- More observations at the same time

3) Partial result when the latent Gamma process is not discrete.

1. Random Measures and their simulation

2. Bayesian inference on Poisson-Gamma RMs

3. An application: inference on Forest Fires in the Amazon Forest

1. **Random Measures and their simulation**
2. Bayesian inference on Poisson-Gamma RMs
3. An application: inference on Forest Fires in the Amazon Forest

Poisson RM and simulation

Definition 1.1.2. A Poisson point process (or Poisson random measure) on $(E, \mathcal{E})$ with intensity measure μ ($\sim ppp(\mu)$) is a point process M such that:

1. for all $B_1, \ldots, B_k \subset E$ measurable and disjoint, $M(B_1), \ldots, M(B_k)$ are independent;
2. for all $B \subset E$ measurable, $M(B) \sim Poisson(\mu(B))$.

Laplace Functional:

$$\boxed{L_M(u) = \exp\left(-\int_E \left(1 - e^{-u(x)}\right) \mu(dx)\right)}$$

Simulation ($\mu(E) < \infty$):

Algorithm 1.1 Simulation of $M \sim ppp(\mu)$

1: Evaluate $A = \mu(E)$

2: Draw one sample z from $Z \sim Poisson(A)$

3: Draw z independent samples from $X_i \sim \frac{\mu(\cdot)}{\mu(E)}$

CRMs

Definition 1.2.1. A Completely Random Measure (CRM) on a Polish space $(E, \mathcal{E})$ is a random measure M such that, for all $B_1, \dots, B_k \subset E$ measurable and disjoint, $M(B_1), \dots, M(B_k)$ are independent.

Levy measure

Laplace Functional:

$$\boxed{L_M(u) = \exp\left(-\int_{\mathbb{R}_{>0}\times E} \left(1 - e^{-vu(x)}\right) \mu(dvdx)\right). \qquad (1.2)}$$

Definition 1.2.2. A Homogeneous completely random measure of parameter (ν, H), for short hCRM(ν, H), is a CRM with Laplace functional (1.2) in the special case:

$$\mu(dvdx) = \nu(dv)H(dx)$$

Gamma Process

Definition 1.2.3. A Gamma random measure (or Gamma process) W on E with base measure H and rate $\beta > 0$, written $\mathcal{G}aP(H, \beta)$, is a hCRM with Levy intensity $\nu(dv) = v^{-1}e^{-\beta v}dv$. In this case the Levy measure is $\mu(dvdx) = v^{-1}e^{-\beta v}dvH(dx)$ and the Laplace functional (1.2) reduces to:

$$W(A) = \int_{\mathbb{R}_{>0}\times A} vN(dvdx)$$

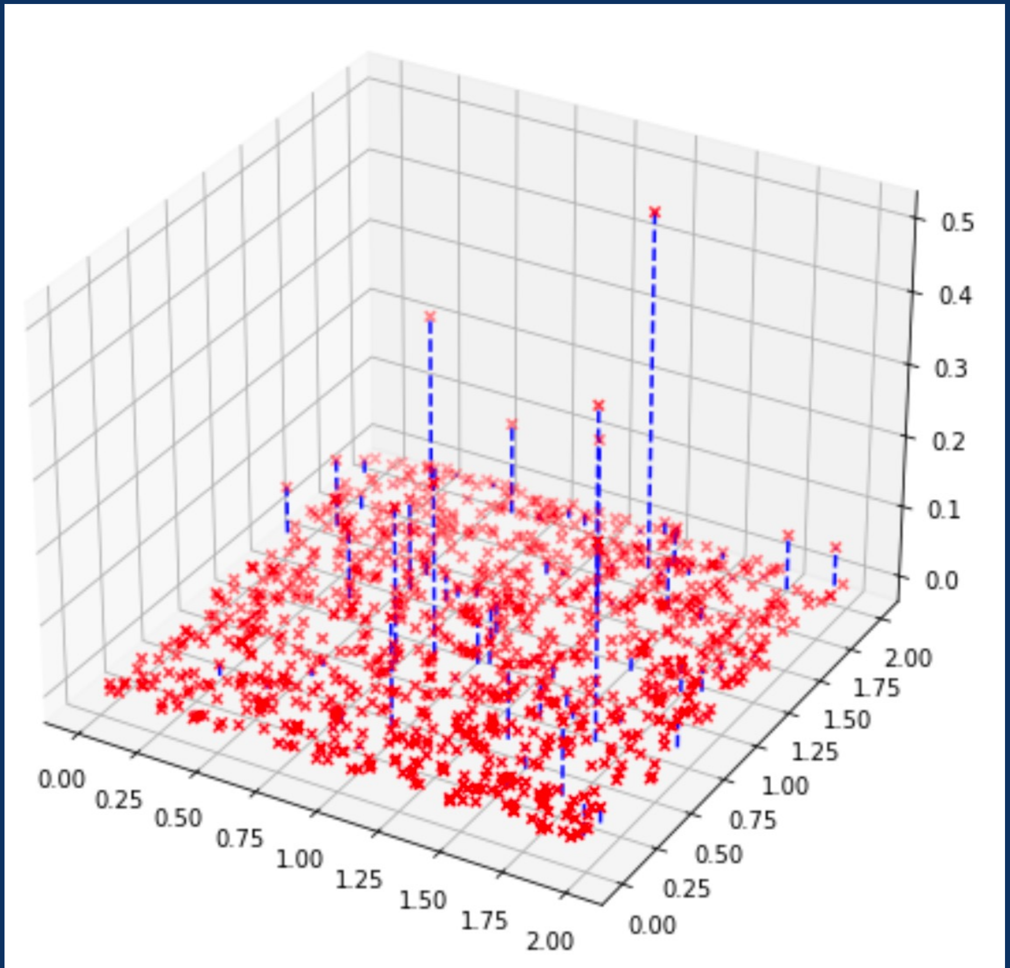
$$= \sum_{j\geq 1} v_j \delta_{x_j}(A)$$

$$L_W(f) = \exp\left(-\int_E \log\left(1 + \frac{f(x)}{\beta}\right) H(dx)\right)$$

Generalised Gamma Process

$$L_W(f) = \exp\left(-\int_E \log\left(1 + \frac{f(x)}{\beta(x)}\right) H(dx)\right)$$

$E = [0, 2] \times [0, 2]$
$M = 1000$
$\Pi(dx) = 1/4\, dx$
$\alpha(x) = 10||x||_1$
$\beta(x) = 2 + x_1$



Simulation of a Gamma Process

Algorithm 1.2 Inverse Levy measure algorithm

1: Initialisation: Fix a large integer M and choose $\Pi(dx)$, a convenient distribution on E, such that $\alpha(x) = \frac{\alpha(dx)}{\Pi(dx)}$

2: Generate M independent identically distributed draws $\{\sigma_m\}_{m \le M}$ from $\Pi(dx)$.

3: Generate $\{\tau_m\}_{m \le M}$, the first M jumps of a standard Poisson process on the real line.

4: Set $v_m = E_1^{-1}\left[\frac{\tau_m}{\alpha(\sigma_m)}\right]\beta(\sigma_m)^{-1}$.

5: Set $\Gamma(dx) \approx \Gamma_M(dx) = \sum_{m \le M} v_m \delta_{\sigma_m}(dx)$.

$$E_1(x) = \lim_{a \to 0}\left[\int_x^{\infty} e^{-t} t^{a-1} dt\right] \qquad E_1^{-1}(x) = \lim_{d \to 0} \frac{1}{2} PPF_d(1 - (d/2)x)$$

[4] R. L. Wolpert, K. Ickstadt. *Poisson/Gamma random field models for spatial statistics.* Biometrika 85(2):251–267, 1998.

Cox Process

Ingredients for a Cox Process:

Point process *W* on *E* (shot noise term)

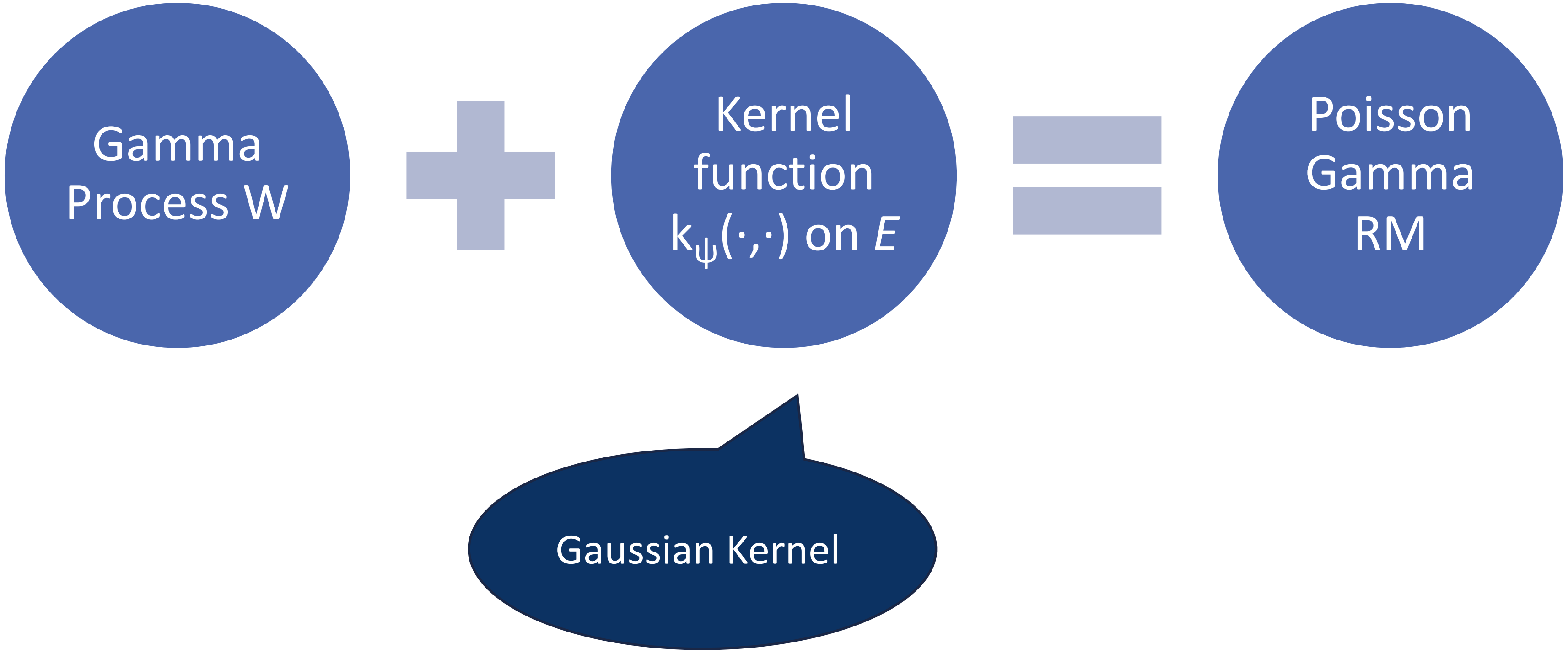
Kernel function $k_\psi(\cdot,\cdot)$ on *E*

$\Lambda(x) = \int_E k_\psi(x,y)W(dy)$ and $\Lambda(dx) = \Lambda(x)w(dx)$

Cox Process = "doubly-stochastic Poisson Process": $\boxed{N(dx)|\Lambda \sim ppp(\Lambda(dx))}$

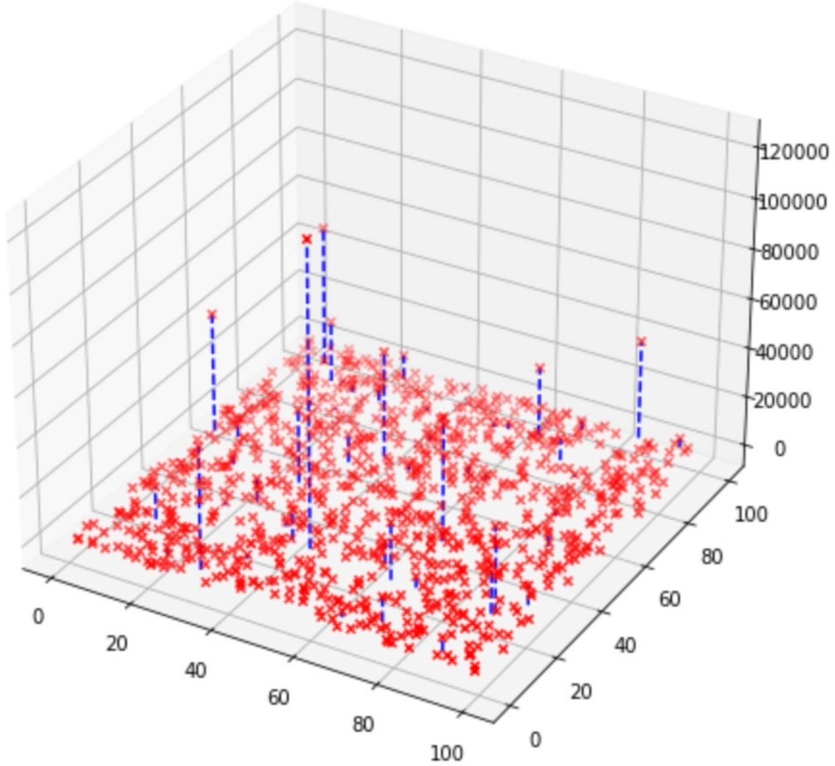
Poisson-Gamma RM

Definition 1.4.2. A Poisson-Gamma random measure is a shot noise Cox process $N(dx)$, where W is a Gamma process, and $\Lambda(dx)$ is derived as above from $\Lambda(x) = \int_E k_\psi(x, y) W(dy)$ for some integral kernel $k_\psi(\cdot, \cdot)$.



Simulation of a Poisson-Gamma RM

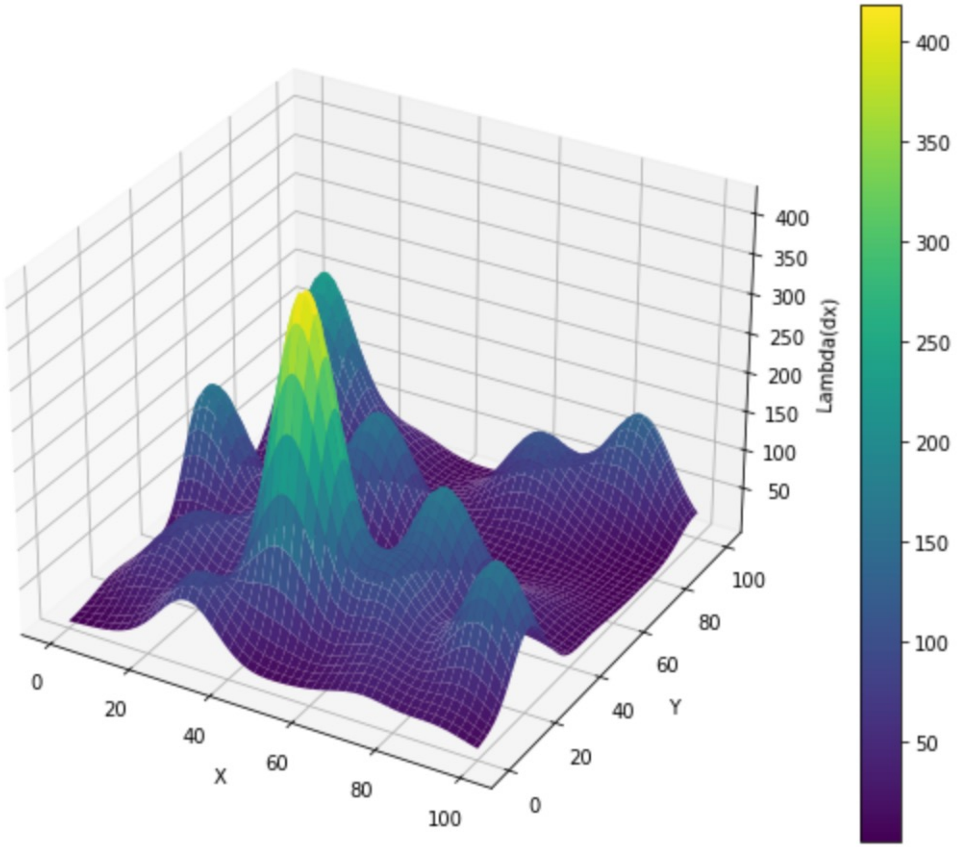
Example 1.4.1. Setting $E = [0, 100] \times [0, 100]$, $M = 1000$, $\beta^{-1} = 1.5 * 10^5$, $H(dx) = \alpha(dx) = 70 * \beta$;



Sample of the underlying Gamma Process W with Algorithm 1.2

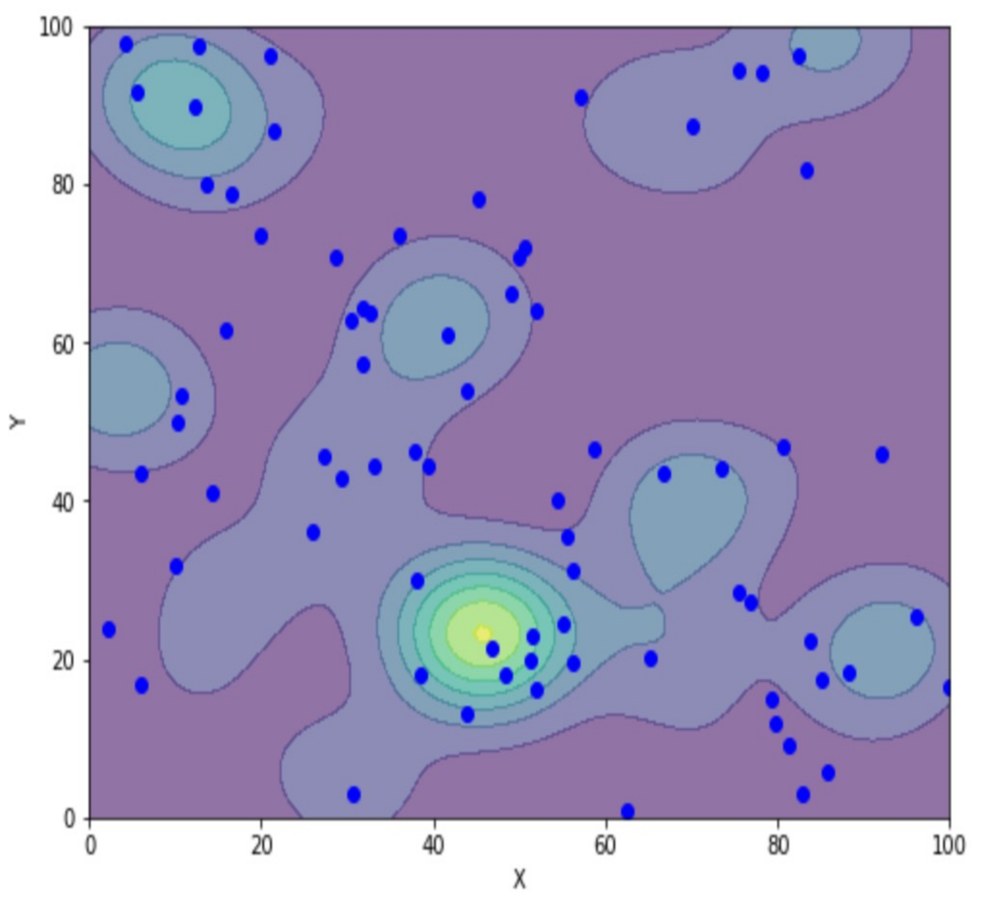
$k_\psi(x, y)$: Gaussian Kernel

$\psi = 5 \quad w(dx) = 10^{-4} dx$



$\Lambda(x)w(dx)$

Points of the Poisson-Gamma process N(dx) on E over the contour plot of Λ(x).



Simulation of a Poisson-Gamma RM

Remember that: $N(dx)|\Lambda \sim ppp(\Lambda(dx))$

$$\Lambda(E) = \sum_{m \leq M} \underbrace{\left\{ v_m \left[\Phi\left(\frac{b-\sigma_m^x}{\psi}\right) - \Phi\left(\frac{a-\sigma_m^x}{\psi}\right) \right] \left[\Phi\left(\frac{d-\sigma_m^y}{\psi}\right) - \Phi\left(\frac{c-\sigma_m^y}{\psi}\right) \right] \right\}}_{w_m}$$

where $\{v_m\}_{m\leq M}$ are the weights and $\{\sigma_m\}_{m\leq M} = \{(\sigma_m^x, \sigma_m^y)\}_{m\leq M}$ are the locations of latent the Gamma process.

Algorithm 1.3 Sampling from $\frac{\Lambda(\cdot)}{\Lambda(E)}$

1: Sample $i \leq M$ with probability $p_m = \frac{w_m}{\Lambda(E)}$
2: **while** $x = (x_i, y_i) \notin E$ **do**
3: Sample $x_i \sim \mathcal{N}(\sigma_i^x, \psi^2)$
4: Sample $y_i \sim \mathcal{N}(\sigma_i^y, \psi^2)$
5: **end while**

1. Random Measures and their simulation

2. **Bayesian inference on Poisson-Gamma RMs**

3. An application: inference on Forest Fires in the Amazon Forest

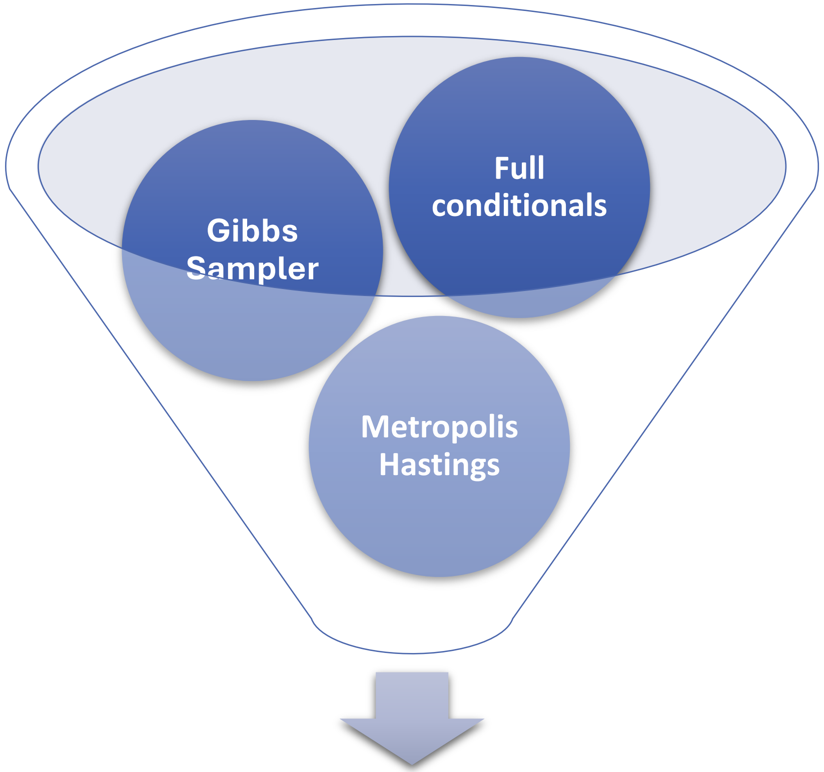
Hierarchical Model for Poisson-Gamma RM:

1. $$\alpha, \beta \sim \pi(\alpha, \beta)$$
2. $$W(dx)|\alpha, \beta \sim \mathcal{G}aP(\alpha(dx), \beta)$$
3. $$\Lambda(dx) = \int_E k_\psi(dx, y) W(dy)$$

$$N(dx)|\alpha, \beta, W \sim ppp(\Lambda(dx))$$

Sampling through MCMC

Objective: Sampling from the posterior of α, β observing $N(dx)$ on E



Metropolis within Gibbs

Definition 3.1.1. A Gamma process on a space E ($\sim \mathcal{G}aP(H_D, \beta)$) is a discrete Gamma process if $H_D(dx)$ is a discrete measure on E.

Proposition 3.1.1. A discrete Gamma process $W(dx) \sim \mathcal{G}aP(H_D, \beta)$ on a space E where $H_D(dx) = \sum_{m=1}^{M} \alpha_m \delta_{\sigma_m}(dx)$ and $\{\sigma_m\}_{m \leq M}$ are fixed points of E, is equivalent to:

$$W(dx) = \sum_{m=1}^{M} v_m \delta_{\sigma_m}(dx)$$

where $v_m \sim \mathcal{G}a(\alpha_m, \beta)$ $\forall m = 1, \ldots, M$ and $v_1, \ldots, v_M$ are independent.

Choice of α

Common sense choice: $\alpha(dx) = \alpha \cdot m(dx)$ $\Longrightarrow$ $\alpha_D(dx) = \sum_{m=1}^{M} \frac{\alpha}{M}\delta_{\sigma_m}(dx)$

$$\left.\begin{array}{l} \alpha_D(x) \longrightarrow \alpha \cdot m(dx) \text{ weakly,} \\ \\ W_D \sim \mathcal{G}aP(\alpha_D(dx), \beta) \\ W \sim \mathcal{G}aP(\alpha \cdot m(dx), \beta) \end{array}\right\} \quad \mathbb{E}\left[e^{-tW_D(A)}\right] \longrightarrow \mathbb{E}\left[e^{-tW(A)}\right]$$

More in Chapter 14 of: O. Kallenberg. *Foundations of Modern Probability*. Springer, 2021.

Properties of the discrete Gamma Process

$$v_1, \dots, v_M \overset{\text{IND}}{\sim} \mathcal{G}a(\alpha_m, \beta) \quad \forall m = 1, \dots, M$$

$$\bar{v} := \sum_{m=1}^{M} v_m \;, \qquad D_m := \frac{v_m}{\bar{v}} \quad \forall m = 1, \dots, M$$

- $\bar{v}$ and $(D_1, \dots, D_M)$ are independent,
- $\bar{v} \sim \mathcal{G}a\left(\sum_{m=1}^{M} \alpha_m, \beta\right)$
- $(D_1, \dots, D_M) \sim DIR(\alpha_1, \dots, \alpha_M)$,

Therefore:

$$W(dx) = \sum_{m=1}^{M} v_m \delta_{\sigma_m}(dx) = \bar{v} \sum_{m=1}^{M} D_m \delta_{\sigma_m}(dx)$$

Likelihood of the model

$$\mathcal{L}(\underline{X}|\alpha_1,\ldots,\alpha_M,\beta) =$$
$$f_{\overline{v}}(\overline{v})f_D(D1,\ldots,D_M)f_N(N|\Lambda(E))$$
$$\times f_{X,S}(X_1,\ldots,X_N,S_1,\ldots,S_N|N,D)$$

$$\alpha = \sum_{m=1}^{M}\alpha_m$$

$$f_{\overline{v}}(\overline{v}) = \frac{\overline{v}^{\alpha-1}e^{-\beta\overline{v}}\beta^{\alpha}}{\Gamma(\alpha)}$$

$$f_D(D1,\ldots,D_M) =$$
$$= \Gamma(\alpha)\cdot\prod_{m=1}^{M}\frac{D_m^{\alpha_m-1}}{\Gamma(\alpha_m)}$$

$$\Lambda(E) = \overline{v}\sum_{m=1}^{M}D_m k_\psi(E|\sigma_m)$$

$$f_N(N|\Lambda(E)) = \frac{\Lambda(E)^N e^{-\Lambda(E)}}{N!}\mathbb{I}_{\mathbb{N}}(N)$$

$$f_{X,S}(X_1,\ldots,X_N,S_1,\ldots,S_N|N,D) =$$
$$= \prod_{i=1}^{N}\frac{D_i k_\psi(E|S_i)}{\sum_{m=1}^{M}D_m k_\psi(E|\sigma_m)}\frac{k_\psi(X_i|S_i)}{k_\psi(E|S_i)}$$

Likelihood of the model

Therefore, supposing (α_1, ..., α_M) and β independent, simplifying and rearranging the terms, the joint likelihood becomes :

$$\mathcal{L}(\underline{X}|\alpha_1,\ldots,\alpha_M,\beta)\pi(\alpha_1,\ldots,\alpha_M,\beta) = \pi(\alpha_1,\ldots,\alpha_M)\pi(\beta)\overline{v}^{\alpha-1}\overline{v}^{N}e^{-\beta\overline{v}}$$
$$\times \exp\left\{-\overline{v}\sum_{m=1}^{M} D_m k_\psi(E|\sigma_m)\right\} \times \beta^{\alpha}\frac{1}{N!}\prod_{m=1}^{M}\frac{D_m^{\alpha_m-1}}{\Gamma(\alpha_m)} \times \prod_{i=1}^{N} D_i k_\psi(X_i|S_i).$$

Likelihood of the model with more observation (time independent)

$$\mathcal{L}(\underline{X_1}, \ldots, \underline{X_T} | \alpha_1, \ldots, \alpha_M, \beta)\pi(\alpha_1, \ldots, \alpha_M, \beta) = \pi(\alpha_1, \ldots, \alpha_M)\pi(\beta)\left(\prod_{t=1}^{T} \overline{v}_t\right)^{\alpha-1}$$

$$\times \prod_{t=1}^{T} \overline{v_t}^{N_t} \times \exp\left(-\beta \sum_{t=1}^{T} \overline{v_t}\right) \times \prod_{t=1}^{T} \exp\left(-\overline{v_t} \sum_{m=1}^{M} D_{m,t} k_\psi(E|\sigma_m)\right)$$

$$\times \beta^{\alpha T} \times \prod_{t=1}^{T} \frac{1}{N_t!} \times \prod_{t=1}^{T} \prod_{m=1}^{M} \frac{D_{m,t}^{\alpha_m - 1}}{\Gamma(\alpha_m)} \times \prod_{t=1}^{T} \prod_{i=1}^{N_t} D_{i,t} k_\psi(X_{i,t}|S_{i,t}).$$

Metropolis within Gibbs Sampler for discrete Poisson-Gamma RMs (1)

Algorithm 3.1 Metropolis within Gibbs Sampler for discrete Poisson-Gamma random measures

1: Initialisation:
- choose arbitrary starting values for: $\overline{v}_t^{(0)}, D_{m,t}^{(0)} \quad \forall t = 1, \dots, T \quad \forall m = 1, \dots, M$;
- choose arbitrary starting values for $\alpha^{(0)}, \beta^{(0)}$;
- choose wisely $\pi(\alpha)$, a prior for α, σ^2_{MC}, the variance for the proposal of α_{new}, and a_0, b_0 small $\in \mathbb{R}$, shape and rate for the prior $\pi(\beta) \sim \mathcal{G}a(a_0, b_0)$.

2: **for** $it \geq 1$ **do**

3: **for** $t = 1, \dots, T$ **do**

4: Assign $S_{i,t}^{(it)}$ with probability:

$$\mathbb{P}(S_{i,t}^{(it)} = \sigma_m) = \frac{D_{i,t}^{(it-1)} k_\psi(X_i|S_i)}{\sum_{m=1}^M D_{m,t}^{(it-1)} k_\psi(X_i|\sigma_m)} \quad \begin{array}{l} \forall i = 1, \dots, N_t \\ \forall m = 1, \dots, M. \end{array}$$

5: Sample: $D_{new,t} = (D_{new,t,1}, \dots, D_{new,M,t}) \sim DIR(\alpha_1^{(it-1)} + n_{1,t}, \dots, \alpha_M^{(it-1)} + n_{M,t})$ and set $D_t^{(it)} = D_{new,t}$ with probability P_1, otherwise $D_t^{(it)} = D_t^{(it-1)}$, where:

$$P_1(D_{new,t}, D_t^{(it-1)}) = \min\left(1, \frac{e^{-\Lambda_{D_{new,t}}(E)}}{e^{-\Lambda_{D_t^{(it-1)}}(E)}}\right).$$

Important choice of hyperparameters

Independence sampler

Metropolis within Gibbs Sampler for discrete Poisson-Gamma RMs (2)

Sample: $\overline{v}_t^{(it)} \sim \mathcal{G}a\left(\alpha^{(it-1)} + N_t, \beta^{(it-1)} + \sum_{m=1}^{M} D_{m,t}^{(it)} k_{\psi}(E|\sigma_m)\right)$.

end for

Sample: $\beta^{(it)} \sim \mathcal{G}a\left(a_0 + \alpha^{(it-1)}T, b_0 + \sum_{t=1}^{T} \overline{v}_t^{(it)}\right)$.

Conjugate model in Beta

Sample: $\alpha_{new} \sim \mathcal{N}(\alpha^{(it-1)}, \sigma^2_{MC})$

and set $\alpha^{(it)} = \alpha_{new}$ with probability P_2, otherwise $\alpha^{(it)} = \alpha^{(it-1)}$, where:

$$\log P_2(\alpha_{new}, \alpha^{(it-1)}) = \min\left(0, \operatorname{logt}(\alpha_{new}) - \operatorname{logt}(\alpha^{(it-1)})\right)$$

and:

$$\operatorname{logt}(\alpha) = \log(\pi(\alpha)) + (\alpha - 1)\sum_{t=1}^{T} \log\left(\overline{v}_t^{(it)}\right) + \alpha T \log\left(\beta^{(it)}\right)$$

$$+\frac{\alpha - M}{M}\sum_{t=1}^{T}\sum_{m=1}^{M} \log\left(D_{m,t}^{(it)}\right) - MT\log\left(\Gamma\left(\frac{\alpha}{M}\right)\right).$$

RW Metropolis

end for

Full conditionals of the model with more Observation (time varying)

If we believe the latent process changes over time, we can capture this aspect allowing α to evolve over time:

Seasonality

Trend over years

$$\alpha_t = e^{\eta_0 + \eta_1 I(t) + \eta_2 A(t)}$$

The full conditional of α becomes:

$$\alpha_1, \dots, \alpha_T | All \propto \pi(\alpha_1, \dots, \alpha_T) \left(\prod_{t=1}^{T} \overline{v}_t^{\alpha_t - 1} \right) \beta^{\sum_{t=1}^{T} \alpha_t} \times \prod_{t=1}^{T} \prod_{m=1}^{M} \frac{D_{m,t}^{\alpha_{t,m}-1}}{\Gamma(\alpha_{t,m})}$$

Metropolis within Gibbs Sampler for discrete Poisson-Gamma RMs (3)

Step 9 of Algorithm 3.1 then becomes:

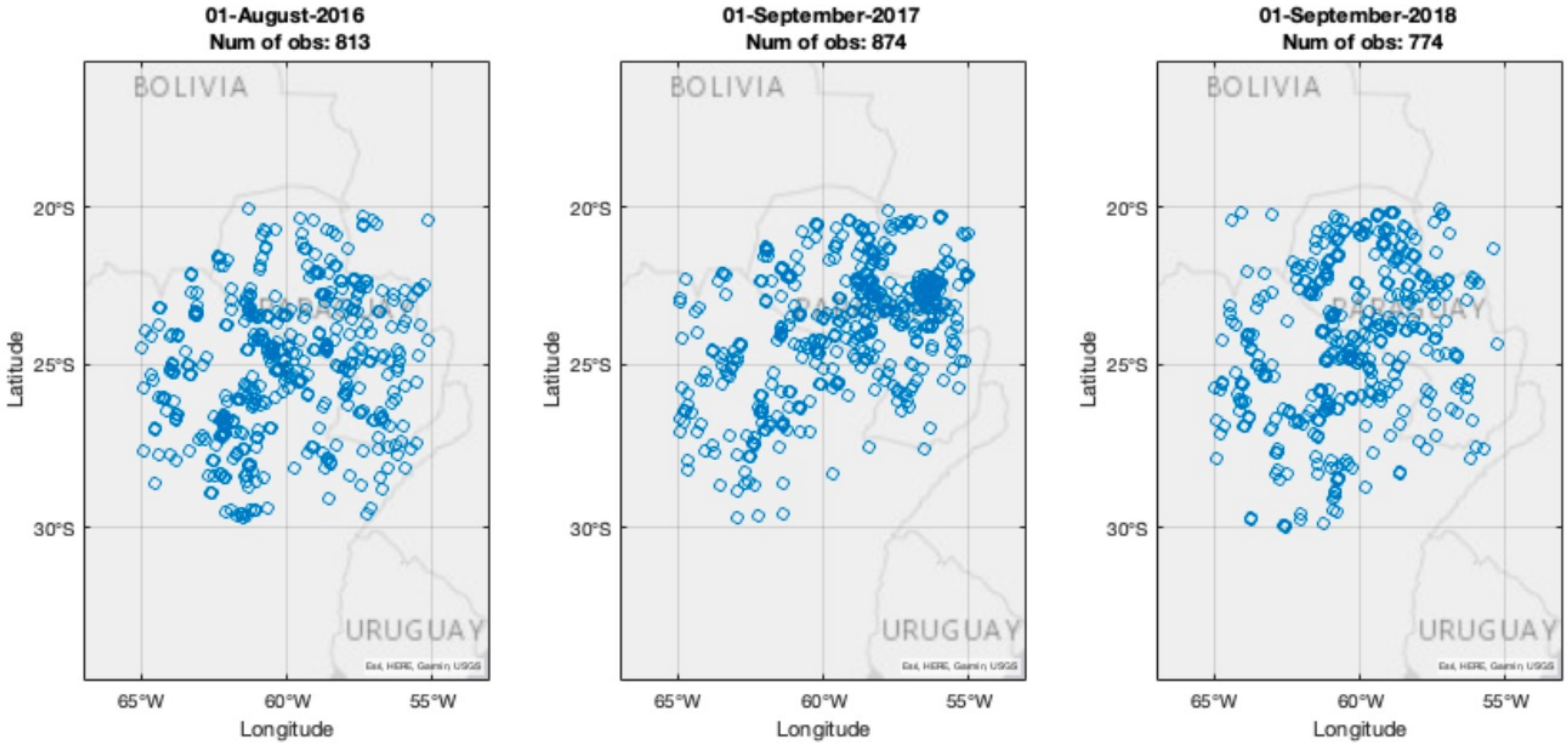
1. Sample: $(\eta_{0,new}, \eta_{1,new}, \eta_{2,new}) \sim \mathcal{N}\left(\left(\eta_0^{(it-1)}, \eta_1^{(it-1)}, \eta_2^{(it-1)}\right), \sigma^2_{MC}\mathbb{I}_3\right)$.
2. Compute: α_{new} and $\alpha^{(it-1)}$ using $\alpha_t = e^{\eta_0+\eta_1 I(t)+\eta_2 A(t)}$.
3. Set $\eta_i^{(it)} = \eta_{i,new} \quad \forall i = 0,1,2$ with probability P_2, otherwise $\eta_i^{(it)} = \eta_i^{(it-1)}$, where:

$$\log P_2(\alpha_{new}, \alpha^{(it-1)}) = \min\left(0, \mathrm{logt}(\alpha_{new}) - \mathrm{logt}(\alpha^{(it)})\right)$$

and: $\mathrm{logt}(\alpha) = \log(\pi(\eta_0)\pi(\eta_1)\pi(\eta_2)) + \sum_{t=1}^{T}(\alpha_t - 1)\log\left(\overline{v}_t^{(it)}\right) + \left(\sum_{t=1}^{T}\alpha_t\right)\log\left(\beta^{(it)}\right)$
$+ \sum_{t=1}^{T}\sum_{m=1}^{M}\left(\frac{\alpha_t}{M} - 1\right)\log\left(D_{m,t}^{(it)}\right) - M\sum_{t=1}^{T}\log\left(\Gamma\left(\frac{\alpha_t}{M}\right)\right)$.

1. Random Measures and their simulation
2. Bayesian inference on Poisson-Gamma RMs
3. **An application: inference on Forest Fires in the Amazon Forest**

The time independent case



Locations of fires as points on a plane based on their coordinates, over three months.

The time independent case

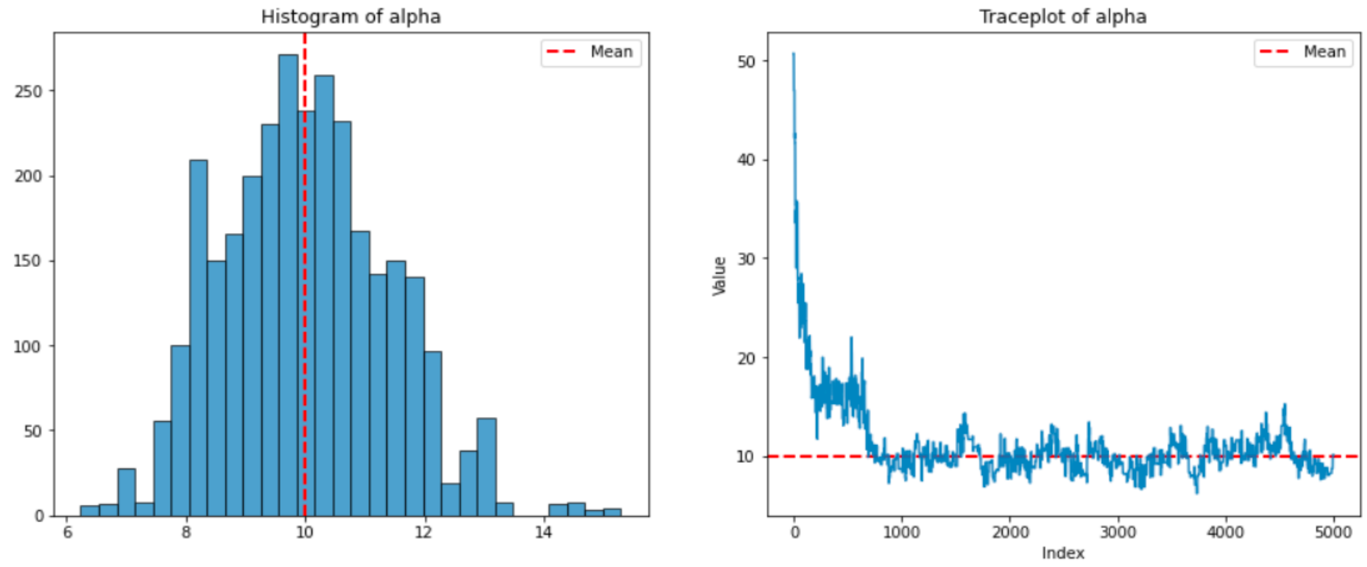


Figure 4.2: Histogram (Burn-in: 2000) and Traceplot of $\alpha^{(it)}$.

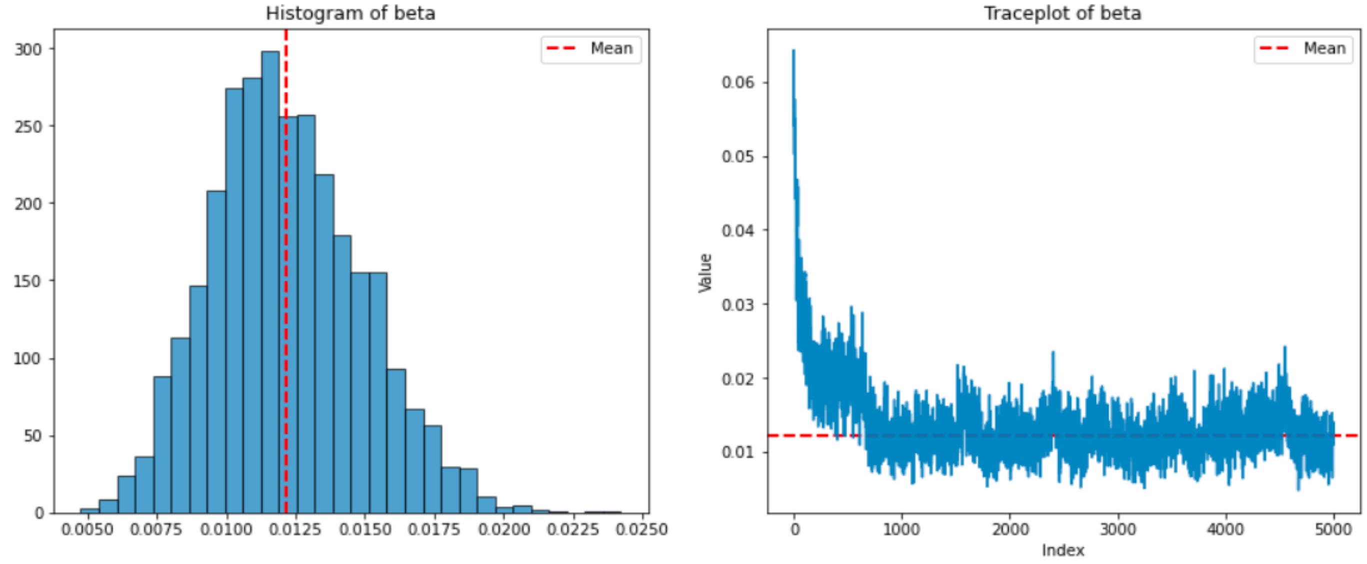


Figure 4.3: Histogram (Burn-in: 2000) and Traceplot of $\beta^{(it)}$.

Running the Sampler for 5000 iterations we obtain samples from the posterior of α and β

The time varying case

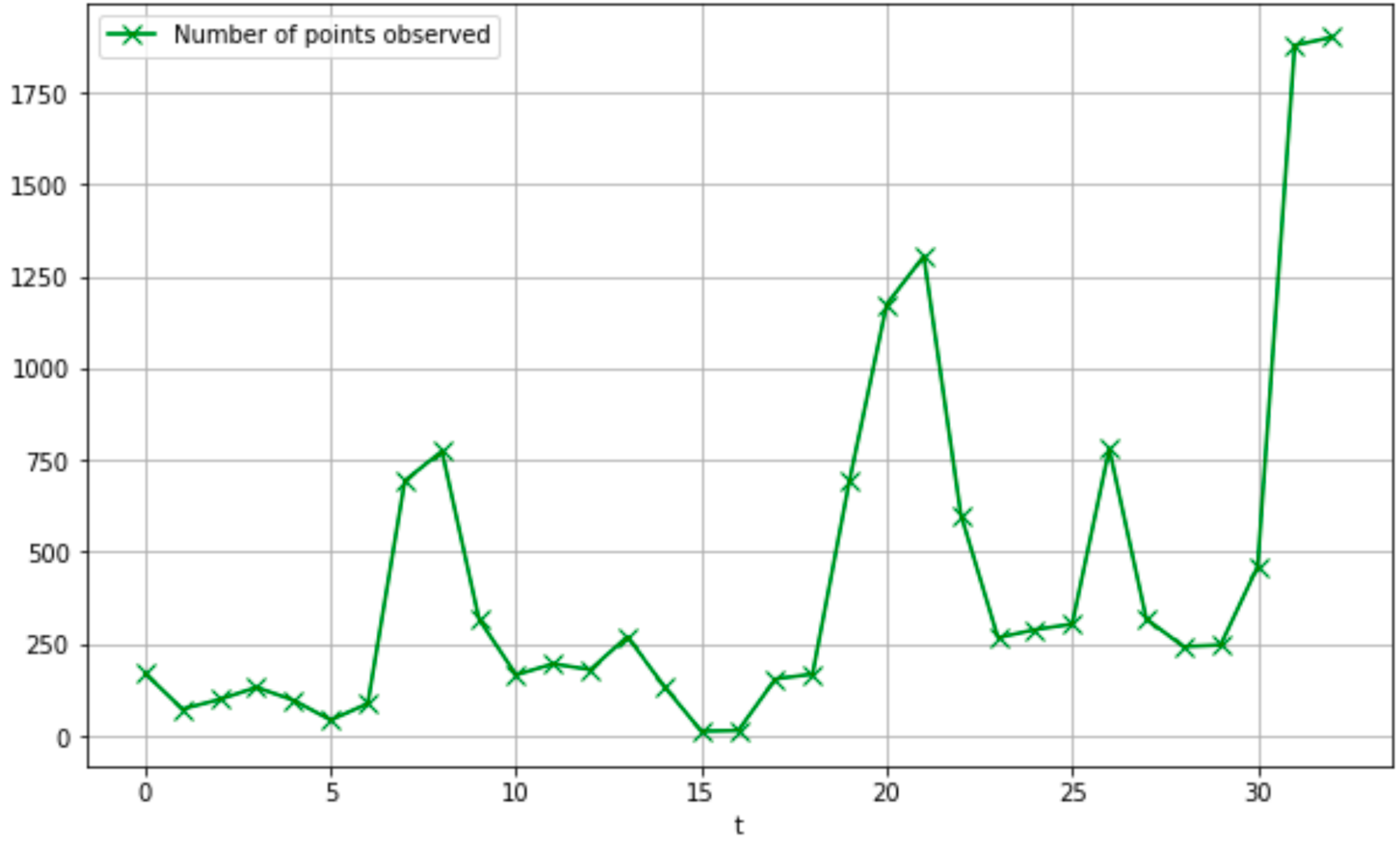


Figure 4.6: Number of forest fires (on the y-axis) observed in the t^{th} month.

Dry vs Wet season (1,0)

$$\alpha_t = e^{\eta_0 + \eta_1 I(t) + \eta_2 A(t)}$$

Year (0,1,2)

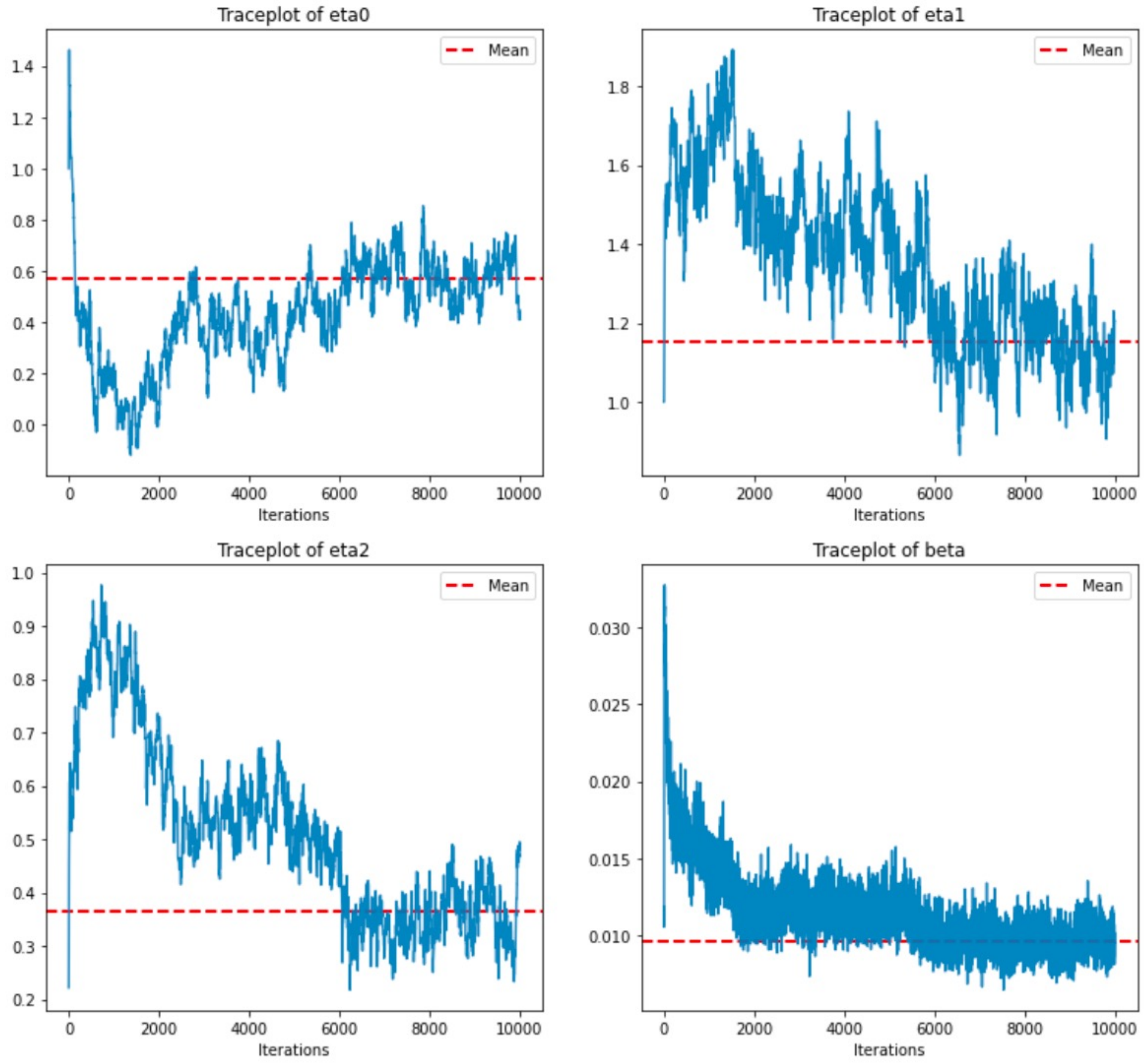
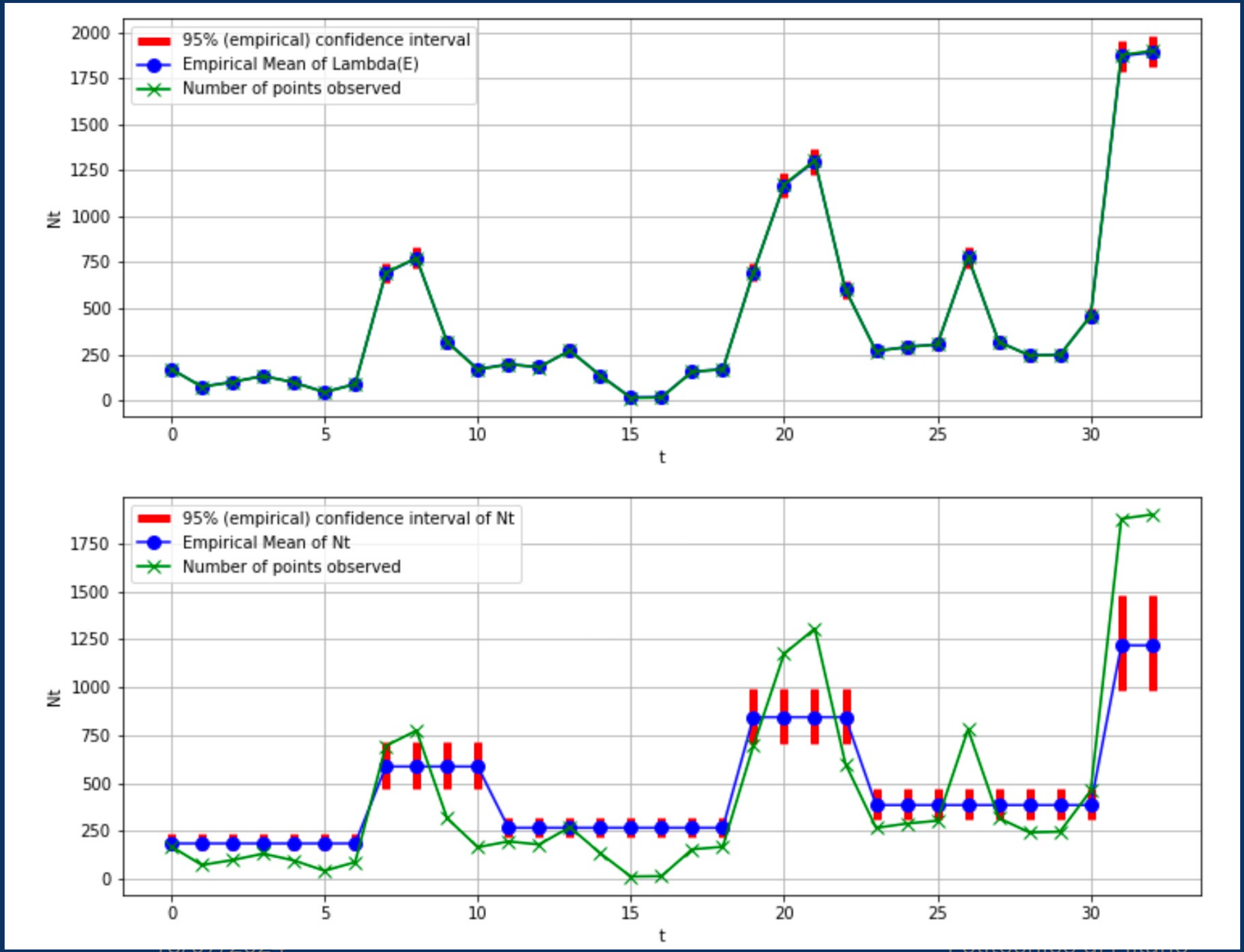


Figure 4.7: Traceplots of $\eta_0, \eta_1, \eta_2, \beta$ and relative means (with a Burn-in of 8000 iterations).

The time varying case

Running the Sampler for 10000 iterations we obtain samples from the posterior of η_0, η_1, η_2 and β.

The time varying case



$$\hat{N}_t = \mathbb{E}(\Lambda_t(E)|\text{Data}) \simeq \frac{1}{L}\sum_{it=1}^{L} \overline{\Lambda_t^{(it)}(E)}$$

$$\tilde{N}_t = \mathbb{E}\left(\frac{\alpha_t}{\beta}|\text{Data}\right) \simeq \frac{1}{L}\sum_{it=1}^{L} \frac{\alpha_t^{(it)}}{\beta^{(it)}}$$

Further work

CONVERGENCE ANALYSIS

ALGORITHMIC REFINEMENT

EXPLORATION OF DIVERSE DATASETS

Bibliography

[1] G. Last, M. Penrose. *Lectures on the Poisson process*. Cambridge University Press, 2018.

[2] D. J. Daley, D. Vere-Jones. *An introduction to the theory of point processes, volume I. Elementary Theory and Methods*. Springer, 2003.

[3] D. J. Daley D. Vere-Jones. *An introduction to the theory of point processes, volume II. General Theory and Structure*. Springer, 2008.

[4] R. L. Wolpert, K. Ickstadt. *Poisson/Gamma random field models for spatial statistics*. Biometrika 85(2):251–267, 1998.

[5] M. Abramowitz, I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. Applied Mathematics Series, 55. Washington, DC: National Bureau of Standards, 1964.

[6] J. M. Marin, C. P. Robert. *A Bayesian Core: A Practical Approach to Computational Bayesian Statistics*. Springer, 2007.

[7] F.Bassetti, R. Casarin, M. Iacopini. *A Spatiotemporal Gamma Shot Noise Cox Process*. ArXiv eprint arXiv:2308.08481, 2023.

[8] O. Kallenberg. *Foundations of Modern Probability*. Springer, 2021.

Grazie per l’Attenzione

Appendix

MCMC (Metropolis-Hastings)

Objective: sampling from a target distribution $\pi(\theta)$

Proposal Kernel: $q(\theta_{new}|\theta_{old})$

Acceptance Probability: $P(\theta_{new}, \theta_{old}) = \min\left(1, \frac{\pi(\theta_{new})q(\theta_{old}|\theta_{new})}{\pi(\theta_{old})q(\theta_{new}|\theta_{old})}\right)$

Algorithm 2.1 Generic Metropolis–Hastings Sampler

1: Initialisation: Choose an arbitrary starting value $\theta^{(0)}$
2: **for** $t \geq 1$ **do**
3: Given $\theta^{(t-1)}$, sample $\hat{\theta}$ from $q\left(\theta|\theta^{(t-1)}\right)$
4: Compute $P = P\left(\hat{\theta}, \theta^{(t-1)}\right)$ as above
5: With probability P accept $\hat{\theta}$ and set $\theta_t = \hat{\theta}$; otherwise, reject $\hat{\theta}$ and set $\theta^{(t)} = \theta^{(t-1)}$
6: **end for**

Independence Sampler: $q(\theta_{new}|\theta_{old}) = q(\theta_{new})$

Random Walk Sampler: $q(\theta_{new}|\theta_{old}) = f(\theta_{new} - \theta_{old})$

MCMC (Gibbs Sampler)

Let $\theta = (\theta_1, \ldots, \theta_p) \in \mathbb{R}^p$ and set $\theta_{-i} = (\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_p)$.

Full conditional: $\pi_i(\theta_i | \theta_{-i}) \quad \forall i = 1, \ldots, p$

Algorithm 2.2 Gibbs Sampler

Initialisation: Choose an arbitrary starting value $\theta^{(0)} = \left(\theta_1^{(0)}, \ldots, \theta_p^{(0)}\right)$

for $t \geq 1$ **do**

Given $\theta^{(t-1)}$, we sample θ^t with the following scheme:

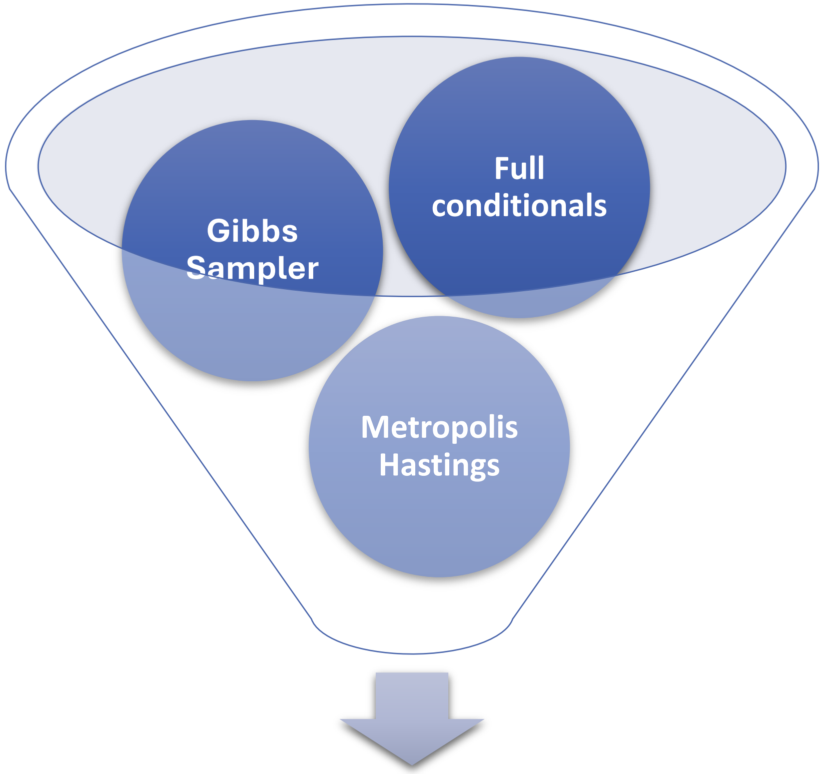
1) $\theta_1^{(t)} \sim \pi_1\left(\theta_1 | \theta_2^{(t-1)}, \ldots, \theta_p^{(t-1)}\right)$

2) $\theta_2^{(t)} \sim \pi_2\left(\theta_2 | \theta_1^{(t)}, \theta_3^{(t-1)}, \ldots, \theta_p^{(t-1)}\right)$

$\ldots$

p) $\theta_p^{(t)} \sim \pi_p\left(\theta_p | \theta_1^{(t)}, \ldots, \theta_{p-1}^{(t)}\right)$

end for



Metropolis within Gibbs

You should be able to sample (in closed form) from the full conditionals!

Proof of Proposition 3.1.1.

Proof. Since the Gamma process W is a CRM, for all $A_1, \dots, A_k \subset E$ measurable and disjoint, $W(A_1), \dots, W(A_k)$ are independent.

Moreover, $\forall A \in E$ and $\forall t > 0$, $tW(A) = \int_E t\mathbb{I}_A(x)W(dx)$.

Therefore:

$$\begin{aligned}
L_{W(A)}(t) &= \mathbb{E}\left[e^{-tW(A)}\right] \\
&= \mathbb{E}\left[e^{-\int_E t\mathbb{I}_A(x)W(dx)}\right] \\
&= \exp\left\{-\int_E \log\left(1 + \frac{t\mathbb{I}_A(x)}{\beta}\right) H_D(dx)\right\} \\
&= \exp\left\{-\log\left(1 + \frac{t}{\beta}\right) H_D(A)\right\} \\
&= \left(1 + \frac{t}{\beta}\right)^{-H_D(A)}
\end{aligned}$$

Which is equivalent to: $W(A) \sim \mathcal{G}a(H_D(A), \beta)$ if $H_D(A) > 0$ or $W(A) = 0$ a.s. if $H_D(A) = 0$.

Proof of Proposition 3.1.1.

Now, since $H_D(dx) = \sum_{m=1}^{M} \alpha_m \delta_{\sigma_m}(dx)$:

$$\Rightarrow H_D(A) = \sum_{m:\sigma_m \in A} \alpha_m$$

$$\Rightarrow W(A) = 0 \text{ a.s.} \quad \text{if } \sigma_m \notin A \quad \forall m = 1, \dots, M$$

$$\Rightarrow W_\omega(A) = \sum_{m=1}^{M} v_m(\omega)\delta_{\sigma_m}(A) \text{ a.s.}$$

which means that $v_m = W(A_m)$ if $\sigma_m \in A_m$ and $\sigma_k \notin A_m$ $\forall k \neq m$.

Therefore, $v_m \sim \mathcal{G}a(\alpha_m, \beta)$ $\forall m = 1, \dots, M$ and $v_1, \dots, v_M$ are independent. $\square$

The time independent case

Figure 4.4: On the left: plot of the coordinates of the fires (blue points) with estimates for the weights v_m (green dashed lines) of the latent process $W(dx)$; On the right: surface plot of $\mathbb{E}(\Lambda|\text{Data})$ on E.

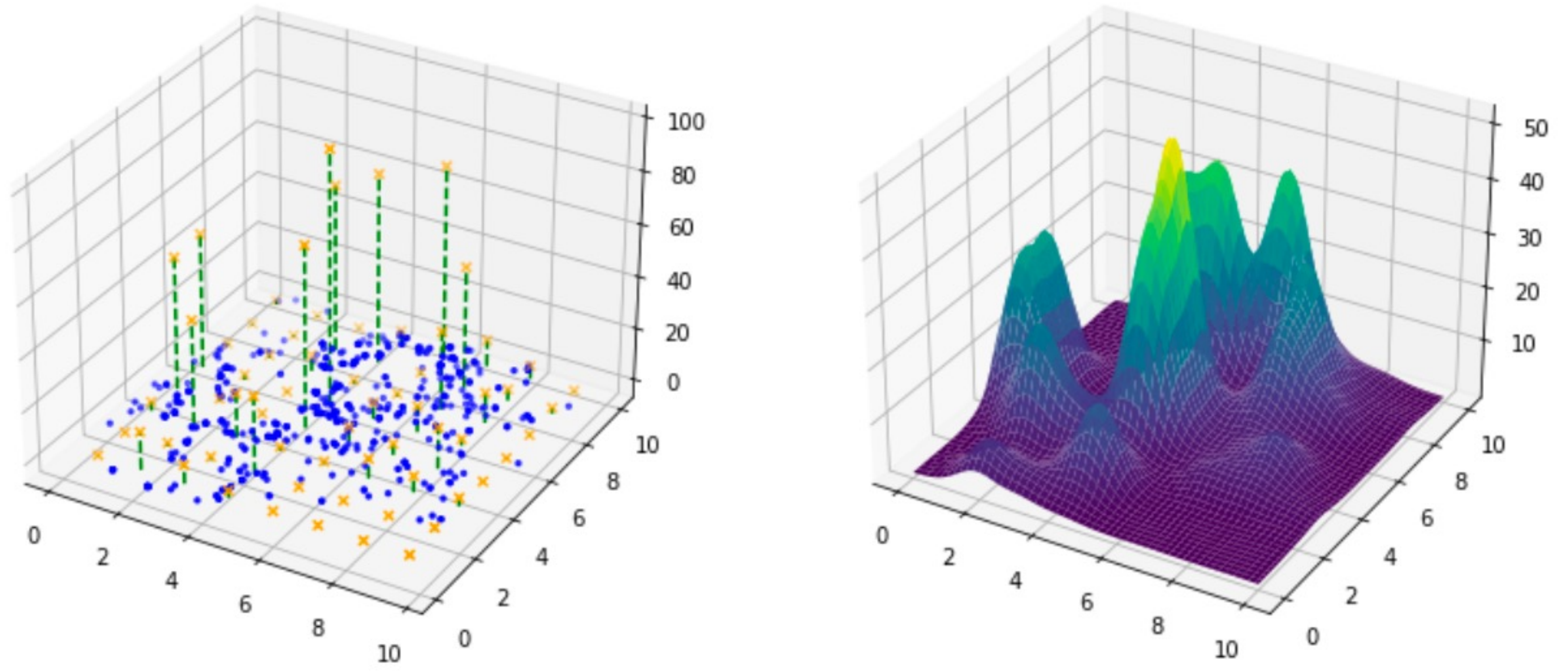


Figure 4.5: Plot of the locations (blue points) of the fires over the contour plot of $\mathbb{E}(\Lambda|\text{Data})$ on E.