POISSON PROCESSES AND 911 CALLS

Jacob Mortensen

March 18, 2015

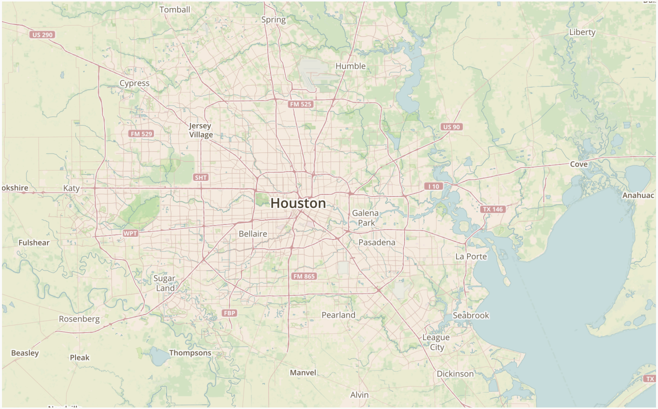
Brigham Young University

INTRODUCTION

INTRODUCTION

- Extreme heat scenarios pose a significant threat to public health
- It is of interest for researchers to understand how the effect of heat varies by location

INTRODUCTION



Our data includes information about 1389 heat-related 911 calls

How many 911 calls will come in and where they will come from?

Spatial data is:

- nonlinear
- highly correlated

so how do we model it?

We use a point process model

- Models both frequency and location of events
- Relies on an intensity function, $\Lambda(x)$
- $\Lambda(x)$ integrates to the expected number of events
- If we normalize $\Lambda(x)$ it becomes the density of event location

$\Lambda(x)$ is the key!

MODEL

MODEL

- $\mathsf{s}_1, \mathsf{s}_2, \ldots, \mathsf{s}_\mathsf{N}$ where N = 1389 are the latitude-longitude coordinates of the observed 911 calls
- $\mathcal{S}$ is the set of all latitude-longitude coordinates in Houston
- $N \sim Pois(\int_{\mathcal{S}} \Lambda(s))$
- $f(\mathsf{s}_1, \mathsf{s}_2, \ldots, \mathsf{s}_\mathsf{N} | N) = \prod_{i=1}^{N} \frac{\Lambda(\mathsf{s}_\mathsf{i})}{\int_{\mathcal{S}} \Lambda(\mathsf{s})}$

Our likelihood is:

$$L(\Lambda(\mathbf{s})) = \frac{\exp(-\int_{\mathcal{S}} \Lambda(\mathbf{s})) \left(\int_{\mathcal{S}} \Lambda(\mathbf{s})\right)^N}{N!} \prod_{i=1}^{N} \frac{\Lambda(\mathbf{s}_\mathrm{i})}{\int_{\mathcal{S}} \Lambda(\mathbf{s})}$$

which is equivalent to:

$$L(\Lambda(\mathbf{s})) = \exp\left(-\int_{\mathcal{S}} \Lambda(\mathbf{s})\right) \prod_{i=1}^{N} \Lambda(\mathbf{s}_i)$$

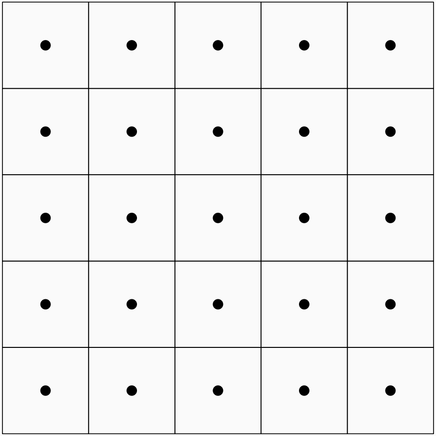
- To simplify we are going to divide $\mathcal{S}$ into a grid
- $\mathsf{s_1}^*, \mathsf{s_2}^*, \ldots, \mathsf{s_K}^*$ where K = 1428 are the latitude-longitude coordinates of the prediction locations
- This allows us to discretize $\Lambda(\mathsf{s})$ to make modeling easier:

$$\Lambda(\mathsf{s}) = \delta \sum_{k=1}^{K} \frac{\lambda_k}{|\mathcal{G}_k|} \mathbb{1}\left\{\mathsf{s} \in \mathcal{G}_k\right\}$$

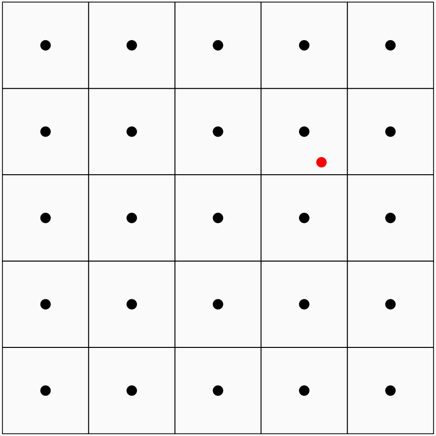
$$\Lambda(\mathbf{s}) = \delta \sum_{k=1}^{K} \frac{\lambda_k}{|\mathcal{G}_k|} \mathbb{1}\left\{\mathbf{s} \in \mathcal{G}_k\right\}$$

- δ is the expected number of 911 calls
- λ_k is the probability that a 911 call occurs in grid cell $\mathcal{G}_k$
- $\mathcal{G}_k$ is the set of spatial coordinates that are closest to grid point $\mathbf{s}_k^*$
- $|\mathcal{G}_k| = \int_{\mathcal{G}_k} d\mathbf{s}$ is the the "area" of grid cell $\mathcal{G}_k$
- $\lambda_k \in (0, 1)$ and $\sum_{k=1}^{K} \lambda_k = 1$

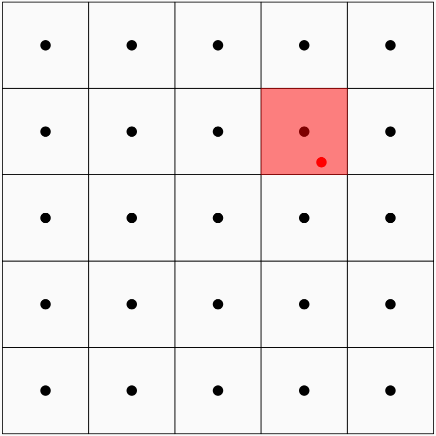
MODEL



MODEL



MODEL



With $\Lambda(\mathsf{s}) = \delta \sum_{k=1}^{K} \frac{\lambda_k}{|\mathcal{G}_k|} \mathbb{1}\left\{\mathsf{s} \in \mathcal{G}_k\right\}$ our likelihood can be written

$$L(\Lambda(\mathsf{s})) = \exp\left(-\delta\right) \delta^N \prod_{i=1}^{N} \frac{\lambda_{\{k:\mathsf{s}_i \in \mathcal{G}_k\}}}{|\mathcal{G}_k|}$$

FITTING THE MODEL

If we use a gamma$(0.001, 0.001)$ distribution as the prior for δ then

$$\pi(\delta)L(\delta|N) \propto \exp(-1.001\delta)\delta^{N+0.001-1}$$

which we recognize as the kernel of a gamma$(N + 0.001, 1.001)$ distribution

How can we use the spatial smoothness in $\Lambda(\mathsf{s})$ to estimate $\lambda_1, \ldots, \lambda_k$?

How can we use the spatial smoothness in $\Lambda(\mathbf{s})$ to estimate $\lambda_1, \ldots, \lambda_k$?

We use a Gaussian process:

- Let $\lambda_k \propto \exp(\lambda_k^*)$
- Let $\boldsymbol{\lambda}^* = \begin{pmatrix} \lambda_1^* \\ \vdots \\ \lambda_k^* \end{pmatrix} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{M})$
- $\mathbf{M}$ is a $K \times K$ correlation matrix with ij^{th} element $M_\nu(\|\mathbf{s}_i^* - \mathbf{s}_j^*\| \,|\phi)$ where $M_\nu(d|\phi)$ is the Matérn correlation function with smoothness ν and decay ϕ
- Then it follows that $\lambda_k = \frac{\exp(\lambda_k^*)}{\sum_{j=1}^{K} \exp(\lambda_j^*)}$

If we use an inverse-gamma(0.01, 0.01) distribution as the prior for σ^2 then

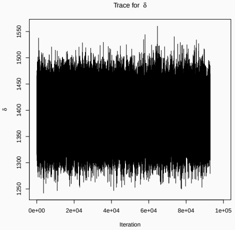
$$\pi(\sigma^2)L(\sigma^2|\boldsymbol{\lambda}^*) \propto (\sigma^2)^{-(0.01+\frac{N}{2})} \exp\left\{-\frac{1}{2\sigma^2}\left[0.01 + \frac{1}{2}(\boldsymbol{\lambda}^*)'\mathsf{M}^{-1}(\boldsymbol{\lambda}^*)\right]\right\}$$

which we recognize as the kernel of an inverse-gamma($0.01 + \frac{N}{2}, 0.01 + \frac{1}{2}(\boldsymbol{\lambda}^*)'\mathsf{M}^{-1}(\boldsymbol{\lambda}^*)$) distribution

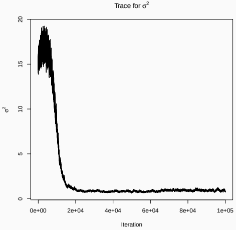
With $\boldsymbol{\lambda}^* \sim \mathcal{N}(0, \sigma^2 \mathsf{M})$,

$$\pi(\boldsymbol{\lambda}^*) L(\boldsymbol{\lambda}^* | N) \propto \exp\left(-\frac{1}{2\sigma^2} (\boldsymbol{\lambda}^*)' \mathsf{M}^{-1} (\boldsymbol{\lambda}^*)\right) \prod_{k=1}^{K} \lambda_k^{N_k}$$

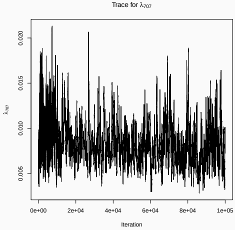
Using a Metropolis within Gibbs sampler, we generate 100,000 draws for $\boldsymbol{\lambda}^*$, δ and σ^2



Trace for δ



Trace for σ^2

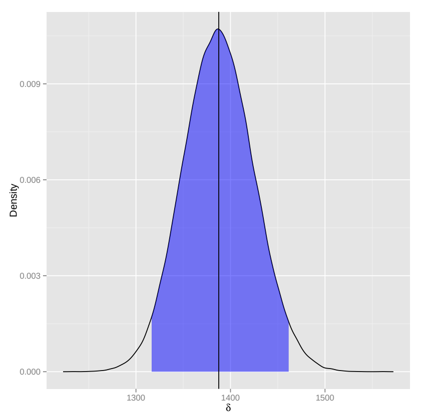


Trace for λ_{707}

RESULTS

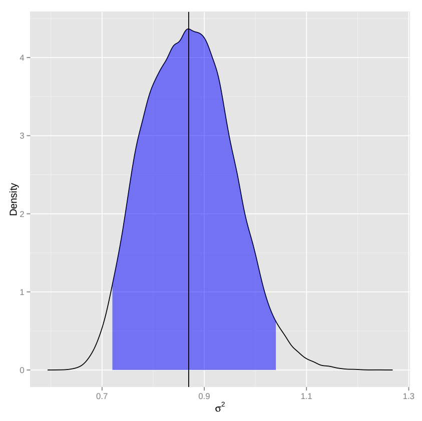
RESULTS

Posterior Density of δ



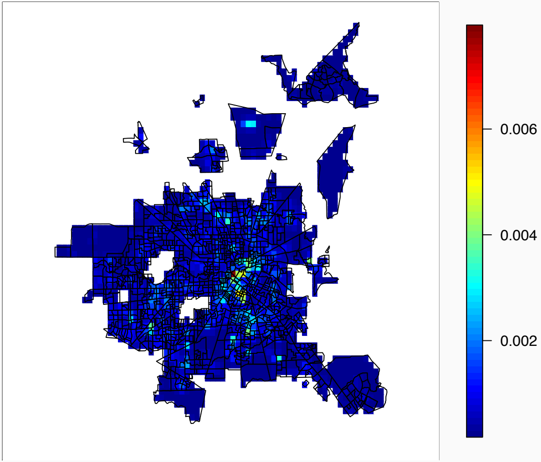
Posterior Mean: 1387.65

Posterior Density of σ^2



Posterior Mean: 0.87

Posterior Density of $\boldsymbol{\lambda}$



QUESTIONS?