Nonparametric Bayesian Spatial Statistics

By Robert L. Wolpert
Duke University, USA

1 Invitation to Bayesian Analysis
2 Example: Rare Poisson Events
3 Hierarchical and Spatial Models
   4 Continuous Models: Remote Sensing
   5 Moving Average Spatial Models
   6 Nonparametric Spatial Epidemiology

3. Hierarchical Models and Bayesian Computation

- I.I.D. Poisson Models
- Exchangeable Poisson Models
- Spatial Poisson Models
- Continuous Spatial Poisson Models
- Example: Biodiversity in Duke Forest

Overstory Trees ($D > 25cm$) in Bormann Plot

Big Oak Trees: How many per $20 \times 20$ m square?

- Guess: 2? 5? 10?
- Density: $f(\lambda) \propto \lambda^\alpha e^{-\lambda}$
- Mean: $E[\lambda] = \alpha / \tau \approx 5$
- Variance: $V[\lambda] = \alpha / \tau^2 \approx 25$
- $\Rightarrow \alpha = 1, \tau = 0.20$:
Now $\lambda \sim \text{Ga}(\alpha, \tau)$ has “prior” density $f(\lambda) \propto \lambda^\alpha e^{-\tau \lambda}$ and, conditional on $\lambda$, $\{x_i\}$ are independent $\text{Po}(\lambda)$; the joint density is then

$$f(\lambda, x_1, \ldots, x_n) = \frac{\tau^\alpha \lambda^\alpha e^{-\tau \lambda}}{\Gamma(\alpha)} \times \prod \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} = e^{\alpha \lambda + \sum x_i - \lambda} e^{-(\tau + \lambda) \lambda}$$

so, **conditional** on $x_1, \ldots, x_n$, the **posterior** distribution of $\lambda$ is $\lambda \sim \text{Ga}(\alpha + \Sigma x_i, \tau + n)$ with mean and variance

$$E[\lambda|x_1, \ldots, x_n] = \frac{\alpha + \sum x_i}{\tau + n} = \frac{\alpha}{\tau} + q \frac{\sum x_i}{n}$$

$$V[\lambda|x_1, \ldots, x_n] = \frac{\alpha + \sum x_i}{(\tau + n)^2}$$

where $p = \frac{\tau}{\tau + n}$, $q = \frac{n}{\tau + n} = 1 - p.$

### Large Oaks: Data from the Bormann Plot

<table>
<thead>
<tr>
<th>$x_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>2</td>
</tr>
</tbody>
</table>

- **Total**: $\Sigma x_i = 162$
- **Mean**: $\bar{x} = 162/25 = 6.48$
- **Variance**: $s^2 = 6.926667 \Rightarrow s = 2.631856$

### Table 1: Oak tree counts in $20 \times 20$ quadrats

### Oaks Distribution, Raw Data:

Oaks: **BUGS** Program for IID Oak Model

```r
model {
  for(i in 1:N) {
    for(j in 1:N) {
      x[i,j] ~ dpois(lambda);
    }
  }
  lambda ~ dgamma(alpha,tau);
}
data:
  list(alpha=1.0, tau=0.20, N=5, M=5,
       x=structure(.Data=c(9,3,8,9,6, 6,4,12,9,2,
                          8,7,10,5, 3,6,9,6,7, 3,8,5,8,2),
                     .Dim=c(5,5)));
init:
  list(lambda=1.0);```
IID Oaks Results:

BUGS reports mean $\mu = 6.466$ with standard deviation $\sigma = 0.5065$ after 100,000 iterations, with an estimated Monte Carlo sampling error of ±0.00166. The exact solution (from above) is

$$E[\lambda|x_1, ..., x_n] = \frac{\alpha + \Sigma x_i}{\tau + n} = \frac{1 + 162}{0.20 + 25} = 6.468254$$

$$V[\lambda|x_1, ..., x_n] = \frac{\alpha + \Sigma x_i}{(\tau + n)^2} = \frac{1 + 162}{(0.20 + 25)^2} = 0.506633^2$$

Evidently MCMC simulation works well for this easy problem.

Exchangeable Oaks Model

The simplest generalization of conjugate Poisson-gamma models that can reflect possible positive association among regions is to allow the Poisson means to be independent gamma variates, drawn from a common distribution about which we will learn from the data. We use the following three-stage hierarchical generalized model:

- **Top:** $\theta \sim \pi(\theta)$
- **Mid:** $\lambda_i \sim \text{Ga}(\alpha^0, \tau^0)$
- **Low:** $X_i \sim \text{Po}(\lambda_i)$

We will choose $\pi(\theta)$ such that $\alpha^0 \sim \text{Ga}(2, 2)$ and $\tau^0 = 0.2$.

Oaks: BUGS Program for Exchangeable Oak Model

```plaintext
model {
  for(i in 1:N) {
    for(j in 1:N) {
      x[i,j] ~ dpois(lam[i,j]);
      lam[i,j] ~ dgamma(alpha,tau);
    }
  }
  alpha ~ dgamma(2,2);
}

data:
  list(tau=0.20, N=5, M=5,
       x=structure(.Data=c(9,3,8,9,6, 6,4,12,9,2,
                       8,7,10,5, 3,6,9,6,7, 3,8,5,8,2), .Dim=c(5,5)));

init:
  list(alpha=1.0,
       lam=structure(.Data=c(1,1,1,1,1, 1,1,1,1,1,
                             1,1,1,1,1, 1,1,1,1,1), .Dim=c(5,5)));
```

Exchangeable Oaks Results:

<table>
<thead>
<tr>
<th></th>
<th>9→8.82</th>
<th>3→3.84</th>
<th>8→8.01</th>
<th>9→8.84</th>
<th>6→6.35</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6→6.34</td>
<td>4→4.67</td>
<td>12→11.34</td>
<td>9→8.84</td>
<td>2→3.01</td>
</tr>
<tr>
<td></td>
<td>8→7.99</td>
<td>7→7.16</td>
<td>7→7.16</td>
<td>10→9.67</td>
<td>5→5.50</td>
</tr>
<tr>
<td></td>
<td>3→3.83</td>
<td>6→6.34</td>
<td>9→8.84</td>
<td>6→6.34</td>
<td>7→7.18</td>
</tr>
<tr>
<td></td>
<td>3→3.85</td>
<td>8→7.99</td>
<td>5→5.50</td>
<td>8→7.99</td>
<td>2→3.00</td>
</tr>
</tbody>
</table>

Table 2: Exchangeable Oak posterior means in 20 × 20 quadrats
Spatial Oaks Model

The simplest generalization of conjugate Poisson-gamma models that can reflect possible positive association for neighboring regions is to allow the Poisson means to be, not independent gamma variates, but linear combinations of a collection of independent gamma variates, with overlapping linear combinations leading to dependence. We use the following three-stage hierarchical generalized model:

- **Top:** \( \theta \sim \pi(\theta) \)
- **Mid:** \( \lambda_i = \sum k_{ij}^\theta \Gamma_j \), \( \Gamma_j \sim \text{Ga}(\alpha^\theta_j, \tau^\theta_j) \)
- **Low:** \( X_i \overset{\text{ind}}{\sim} \text{Po}(\lambda_i) \)

with, for example, \( k_{ij}^\theta = \theta_1 \) for \( i = j \), \( k_{ij}^\theta = \theta_2 \) for \( i \sim j \), and otherwise \( k_{ij}^\theta = 0 \).
Oaks: Posterior Mean of Oak Data

<table>
<thead>
<tr>
<th>Raw</th>
<th>20m Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.12</td>
<td>5.35</td>
</tr>
<tr>
<td>7.45</td>
<td>7.75</td>
</tr>
<tr>
<td>6.08</td>
<td>6.17</td>
</tr>
<tr>
<td>6.90</td>
<td>6.08</td>
</tr>
<tr>
<td>7.80</td>
<td>6.54</td>
</tr>
<tr>
<td>8.64</td>
<td>5.95</td>
</tr>
<tr>
<td>7.82</td>
<td>6.93</td>
</tr>
<tr>
<td>6.55</td>
<td>4.38</td>
</tr>
</tbody>
</table>

Table 3: Posterior Means, Spatial Model

Spatial Oaks Results:

The **mean vectors** and **covariance matrices** of the point counts, conditional on $\theta \in \Theta$, follow by routine computation:

\[
\mathbb{E}^\theta[N_i] = \sum_{j \in J} k_{ij}^\theta \left( \tau_j^\theta \right)^{-1} \alpha_j^\theta
\]

\[
\text{Cov}^\theta[N_i, N_{i'}] = \sum_{j \in J} \left( \delta_{ij}^\theta + k_{ij}^\theta \left( \tau_j^\theta \right)^{-1} \right) k_{ij}^\theta \left( \tau_j^\theta \right)^{-1} \alpha_j^\theta
\]

Unconditionally the $\{N_i\}_{i \in I}$ are distributed as sums of independent negative-binomial random variables.

Note NON-uniform shrinkage
What Level of Aggregation Should We Use?

Refinement:

Poisson random fields are infinitely divisible—so we can refine from a $5 \times 5$ grid to a $100 \times 100$ grid to a $1000 \times 1000$ grid, but of course the calculations get harder. It is better to go all the way to $\infty \times \infty$—and pass to a Poisson Random Field $N(dx)$, and at the same time replace $\{\Gamma_j\}$ with a Gamma Random Field $Ga(ds)$, leading to the Continuous Poisson/Gamma Mixture Model:

Continuous Spatial Oaks:

Parameter: $\theta \sim \pi(\theta)$

Impulses: $\Gamma(ds) \sim Ga(\alpha^\theta(ds), \tau^\theta(s))$

Intensities: $\Lambda(x) \equiv \int_S k^\theta(x, s) \Gamma(ds)$

Point counts: $N(dx) \sim Po(\Lambda(s) m(ds))$
Continuous Spatial Oaks:

Oaks and Hickories:

Continuous Spatial Hickories:

How Much Diversity? One Species or Two?
Biodiversity and Hill’s Number

A traditional measure of the “uniformity” of a probability distribution \(\{p_i\}\) is its Shannon Entropy

\[ H(\hat{p}) \equiv -\sum p_i \log p_i \]

Always positive, this measure is bounded above by \(0 \leq H \leq \log n\) if \(p = \{p_1, \ldots, p_n\}\) is concentrated on \(n\) points— that limit is attained if the \(\{p_i\}\) are all equal (necessarily to \(1/n\)).

Mark Hill’s (1973) index \(H_1\):

\[ 1 \leq H_1 \equiv \exp(H) = \prod \left( \frac{1}{p_i} \right)^{p_i} \leq n \]

“Equivalent Number of Species”— also the limit as \(\alpha \to 1\) of

\[ 1 \leq H_\alpha \equiv \left( \sum p_i^\alpha \right)^{1/(1-\alpha)} \leq n \]
Related papers:


