Unbiased estimating functions for spatial point processes

Rasmus Waagepetersen
Department of Mathematical Sciences
Aalborg University

based on joint work (in progress !) with

Adrian Baddeley, Jean-Francois Coeurjolly, Yongtao Guan,
Abdollah Jalilian and Ege Rubak

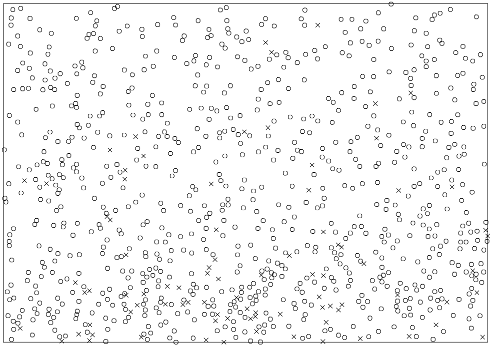
Outline

- Data examples
- GNZ and Campbell formulae
- Gibbs and Cox spatial point processes
- pseudo-likelihood and composite likelihood
- Monte Carlo approximations and relation to logistic regression
- Examples of applications

Aim: discuss closely related estimating functions for two very distinct classes of point processes.

Mucous membrane cells

Centres of cells in mucous membrane:



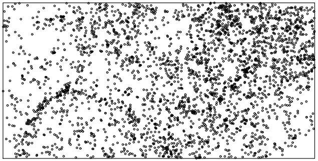
Repulsion due to physical extent of cells

Inhomogeneity - lower intensity in upper part

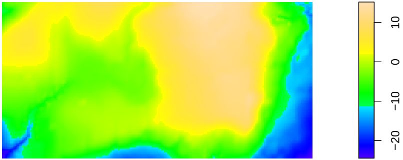
Bivariate - two types of cells

Same type of inhomogeneity for two types ?

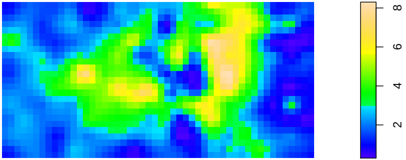
Data example: *Capparis Frondosa*



- *observation window W* $= 1000 \text{ m} \times 500 \text{ m}$
- seed dispersal⇒ *clustering*
- environment ⇒ *inhomogeneity*



Elevation



Potassium content in soil.

Objective: quantify dependence on environmental variables.

Intensity and conditional intensity

Point process $\mathbf{X}$: random point pattern. Assume observed in bounded window $W \subset \mathbb{R}^2$. u spatial location in W.

Intensity $\lambda(u)$: for infinitesimal region A and $u \in A$,

$$P(\mathbf{X} \text{ point in } A) = \lambda(u)|A|$$

Conditional intensity $\lambda(u, \mathbf{X})$:

$$P(\mathbf{X} \text{ has a point in A}|\mathbf{X} \setminus A) = \lambda(u, \mathbf{X})|A|$$

Note

$$P(\mathbf{X} \text{ point in } A) = \mathbb{E}P(\mathbf{X} \text{ point in } A|\mathbf{X} \setminus A) \Rightarrow \lambda(u) = \mathbb{E}\lambda(u, \mathbf{X})$$

GNZ and Campbell formulae

Georgii-Nguyen-Zessin formula:

$$\mathbb{E} \sum_{u \in \mathbf{X}} f(u, \mathbf{X} \setminus u) = \int_W \mathbb{E}[f(u, \mathbf{X}) \lambda(u, \mathbf{X})] \mathrm{d}u$$

for non-negative functions f.

Campbell formula:

$$\mathbb{E} \sum_{u \in \mathbf{X}} f(u) = \int_W f(u) \lambda(u) \mathrm{d}u$$

Note: special case of GNZ since $\lambda(u) = \mathbb{E}\lambda(u, \mathbf{X})$.

Gibbs point processes

Gibbs point processes specified by explicit model for the conditional intensity.

Strauss:

$$\lambda_\theta(u, \mathbf{X}) = \exp[\beta + \psi n_R(u, \mathbf{X})], \quad \beta > 0,\ \psi \le 0$$

$n_R(u, \mathbf{X})$: number of neighboring points within distance R from u.

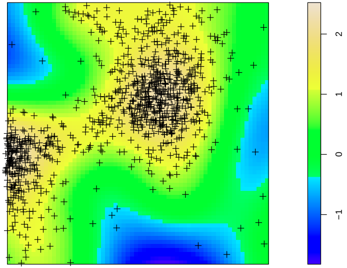
Inhomogeneous: $Z(u)$ covariate at $u \in \mathbb{R}^2$.

$$\lambda_\theta(u, \mathbf{X}) = \exp[\beta Z(u)^{\mathsf{T}} + \psi n_R(u, \mathbf{X})]$$

Cox processes

X Poisson process with intensity function $\lambda(\cdot)$:

total number of points Poisson and given this, points *iid* with density $\propto \lambda(u)$.



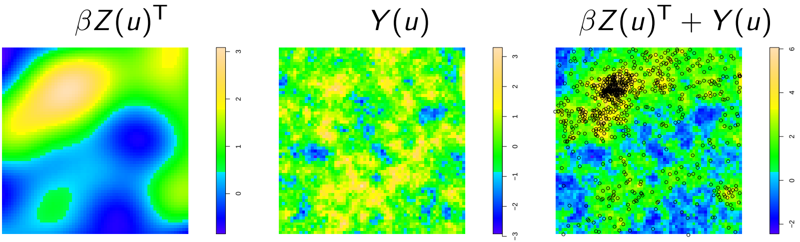
X is a *Cox process* driven by the *random* intensity function Λ if, conditional on $\Lambda = \lambda$, **X** is a Poisson process with intensity function λ.

Example: log Gaussian Cox process

log Gaussian Cox process ("point process GLMM")

$$\Lambda(u) = \exp[\beta Z(u)^{\mathsf{T}} + Y(u)]$$

where $\{Y(u)\}$ Gaussian random field:



For Gibbs point process $\lambda(u, \mathbf{X})$ is given but $\lambda(u) = \mathbb{E}\lambda(u, \mathbf{X})$ hard.

For Cox process, $\lambda(u, \mathbf{X})$ not known but

$$\lambda(u)|A| = P(\mathbf{X} \text{ point in } A) = \mathbb{E}P(\mathbf{X} \text{ point in } A|\Lambda) = \mathbb{E}\Lambda(u)|A|$$

Often $\lambda(u) = \mathbb{E}\Lambda(u)$ easy to evaluate for Cox processes.

E.g. $\log \Lambda(u) \sim N(\beta Z(u)^\mathsf{T}, \sigma^2)$ [log Gaussian Cox process]:

$$\lambda(u) = \exp(\beta Z(u)^\mathsf{T} + \sigma^2/2)$$

Estimating function

Estimating function: $e(\theta)$ $[e(\theta, \mathbf{X})]$ function of θ and data $\mathbf{X}$.

Parameter estimate $\hat{\theta}$ solution of

$$e(\theta) = 0$$

Sensitivity:

$$S = -\mathbb{E}[\frac{\mathrm{d}}{\mathrm{d}\theta} e(\theta)]$$

minus expected derivative of $e(\theta)$

Estimating function

Estimating function: $e(\theta)$ $[e(\theta, \mathbf{X})]$ function of θ and data $\mathbf{X}$.

Parameter estimate $\hat{\theta}$ solution of

$$e(\theta) = 0$$

Sensitivity:

$$S = -\mathbb{E}[\frac{\mathrm{d}}{\mathrm{d}\theta} e(\theta)]$$

minus expected derivative of $e(\theta)$

$\hat{\theta}$ unbiased $\mathbb{E}\hat{\theta} = \theta^*$ if $e(\theta)$ unbiased $\mathbb{E}e(\theta^*) = 0$ (θ^* true value).

$$\mathbb{V}\mathrm{ar}\hat{\theta} = S^{-1}\Sigma S^{-1} \quad \Sigma = \mathbb{V}\mathrm{ar} e(\theta^*)$$

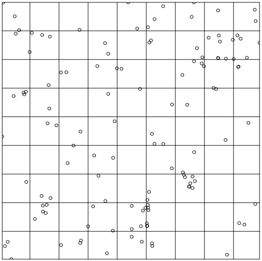
How do we construct unbiased estimating functions involving $\mathbf{X}$ and θ ?

Composite and pseudo-likelihood

Disjoint subdivision $W = \cup_{i=1}^{m} C_i$ in 'cells' C_i.

$u_i \in C_i$ 'center' point.

Random indicator variables:
$N_i = 1[\mathbf{X}$ has a point in $C_i \neq \emptyset]$
(presence/absence of points in C_i).



$$P(N_i = 1) = |C_i|\lambda_\theta(u_i) \text{ and } P(N_i = 1|\mathbf{X} \setminus C_i) = |C_i|\lambda_\theta(u_i, \mathbf{X})$$

Idea: form composite likelihoods based on N_i with marginal or conditional probabilities.

Consider limit when $|C_i| \to 0$.

Log composite likelihood (in fact log likelihood for Poisson):

$$\sum_{u \in \mathbf{X}} \log \lambda_\theta(u) - \int_W \lambda_\theta(u) \mathrm{d}u$$

Log pseudo-likelihood (Besag, 1977)

$$\sum_{u \in \mathbf{X}} \log \lambda_\theta(u, \mathbf{X} \setminus u) - \int_W \lambda_\theta(u, \mathbf{X}) \mathrm{d}u$$

Scores:

$$\sum_{u \in \mathbf{X}} \frac{\lambda'_\theta(u)}{\lambda_\theta(u)} - \int_W \lambda'_\theta(u) \mathrm{d}u$$

and

$$\sum_{u \in \mathbf{X}} \frac{\lambda'_\theta(u, \mathbf{X} \setminus u)}{\lambda_\theta(u, \mathbf{X} \setminus u)} - \int_W \lambda'_\theta(u, \mathbf{X}) \mathrm{d}u$$

unbiased estimating functions by Campbell/GNZ.

Issue:

- integrals

$$\int_W \lambda'_\theta(u)\mathrm{d}u \text{ and } \int_W \lambda'_\theta(u, \mathbf{X})\mathrm{d}u$$

often not explicitly computable.

Numerical quadrature may introduce bias.

Monte Carlo approximation

Let $\mathbf{D}$ 'quadrature/dummy' point process of intensity $\rho(\cdot)$ and independent of $\mathbf{X}$.

By GNZ

$$\mathbb{E} \int_W \lambda'(u, \mathbf{X}) \mathrm{d}u = \mathbb{E} \sum_{u \in \mathbf{X} \cup \mathbf{D}} \frac{\lambda'(u, \mathbf{X})}{\lambda(u, \mathbf{X}) + \rho(u)}$$

By Campbell

$$\int_W \lambda'(u) \mathrm{d}u = \mathbb{E} \sum_{u \in \mathbf{X} \cup \mathbf{D}} \frac{\lambda'(u)}{\lambda(u) + \rho(u)}$$

Idea: replace integrals in pseudo- or composite likelihood with unbiased estimates using $\mathbf{D}$.

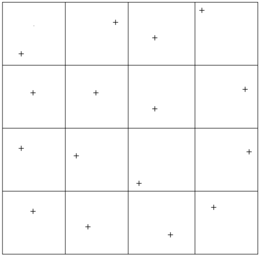
Dummy point process

Should be easy to simulate and mathematically tractable.

Possibilities:

1. Poisson process
2. binomial point process (fixed number of independent points)
3. stratified binomial point process

Stratified:



Monte Carlo approximation and logistic regression

Consider binary variables Y_u with

$$p(u) = P(Y_u = 1) = \frac{f_\theta(u)}{f_\theta(u) + 1}$$

Log logistic regression likelihood:

$$\sum_{u:Y_u=1} \log \frac{f_\theta(u)}{1 + f_\theta(u)} + \sum_{u:Y_u=0} \log \frac{1}{1 + f_\theta(u)} =$$

$$\sum_{u:Y_u=1} \log f_\theta(u) + \sum_{\text{all } u} \log \frac{1}{1 + f_\theta(u)}$$

Score function:

$$\sum_{u:Y_u=1} \frac{f'(u)}{f_\theta(u)} + \sum_{\text{all } u} \frac{f'(u)}{1 + f_\theta(u)}$$

Approximate pseudo- and composite likelihood scores:

$$s(\theta) = \sum_{u \in \mathbf{X}} \frac{\lambda'_\theta(u, \mathbf{X} \setminus u)}{\lambda_\theta(u, \mathbf{X} \setminus u)} - \sum_{u \in (\mathbf{X} \cup \mathbf{D})} \frac{\lambda'_\theta(u, \mathbf{X} \setminus u)}{\lambda_\theta(u, \mathbf{X} \setminus u) + \rho(u)}$$

$$s(\theta) = \sum_{u \in \mathbf{X}} \frac{\lambda'_\theta(u)}{\lambda_\theta(u)} - \sum_{u \in (\mathbf{X} \cup \mathbf{D})} \frac{\lambda'_\theta(u)}{\lambda_\theta(u) + \rho(u)}$$

Note: of logistic regression/case control form with 'probabilities'

$$p(u|\mathbf{X}) = \frac{\lambda_\theta(u, \mathbf{X} \setminus u)}{\lambda_\theta(u, \mathbf{X} \setminus u) + \rho(u)}$$

and

$$p(u) = \frac{\lambda_\theta(u)}{\lambda_\theta(u) + \rho(u)}$$

I.e. probabilities that $u \in \mathbf{X}$ given $u \in \mathbf{X} \cup \mathbf{D}$.

Hence computations straightforward with `glm()` software !

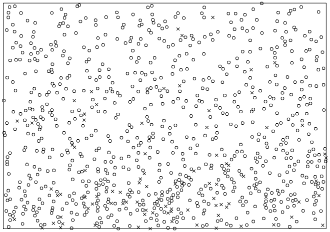
Asymptotic results

Available - but quite technical (will skip details here).

Asymptotic covariance matrix implemented in spatstat $\Rightarrow$ approximate confidence intervals.

Possible to evaluate the proportion of estimation variance due to random quadrature points.

Example: mucous membrane



86 (type 1) + 807 (type 2) points.

1×0.7 observation window.

Marked point $u = (x, y, m)$ where $m = 1$ or 2 (two types of points).

Bivariate Strauss point process with

$$\lambda_\theta(u, \mathbf{X}) = \exp[q_m(y) + \psi n_R(u, \mathbf{X})]$$

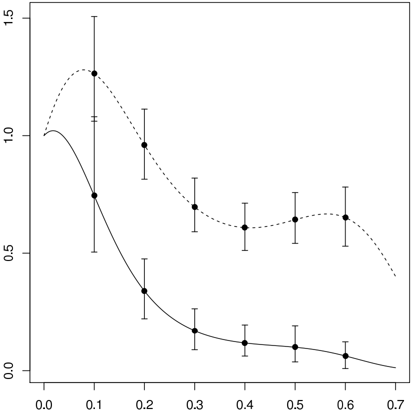
$q_m(y)$: polynomial in spatial y-coordinate.

$n_R(u, \mathbf{X})$: number of neighbors within range $R = 0.008$.

3600 stratified dummy points (random marks 1 or 2).

Fitted polynomials

Fitted polynomials (with confidence intervals for selected *y* values):



Polynomials significantly different according to logistic likelihood ratio test (parametric bootstrap).

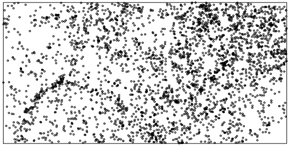
Decomposition of variance

	3600				14400		
	$\hat{\theta}$	sd($\hat{\theta}$)	sd(T_1)	inc. (%)	sd($\hat{\theta}$)	sd(T_1)	inc. (%)
$q_1(0.1)$	6.004	0.195	0.189	3.608	0.191	0.189	0.812
$q_1(0.3)$	4.528	0.267	0.263	1.332	0.264	0.263	0.301
$q_1(0.5)$	3.994	0.406	0.404	0.555	0.404	0.404	0.146
$q_2(0.1)$	7.800	0.091	0.078	15.623	0.082	0.079	3.801
$q_2(0.3)$	7.204	0.083	0.075	10.923	0.076	0.075	2.589
$q_2(0.5)$	7.123	0.086	0.077	10.558	0.080	0.078	2.824
ψ	-2.594	0.344	0.341	0.971	0.342	0.341	0.197

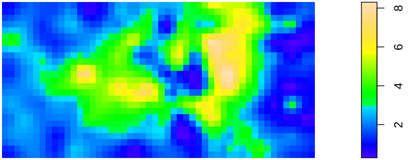
sd(T_1) $\approx$ standard deviation for pseudo-likelihood without approximation.

Example: tree species Capparis Frondosa and Loncocharpus Heptaphyllus

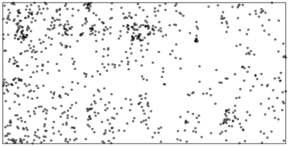
Capparis Frondosa



Potassium content in soil.



Loncocharpus Heptaphyllus



Covariates pH, elevation, gradient, potassium,...

Objective: infer regression model $\lambda_\beta(u) = \exp[\beta Z(u)^\mathsf{T}]$

Clustered point patterns: Cox point process natural model.

Problem: covariates sampled on (coarse) deterministic grid.

Plots shown: interpolated values of covariates.

Hence unbiased Monte Carlo approximation not applicable.

For now: integral

$$\int_W \lambda_\beta(u)\mathrm{d}u$$

approximated using numerical quadrature based on interpolated values.

Need to convince biologists to use random sampling designs.

Optimality ?

Composite likelihood score

$$\sum_{u \in \mathbf{X}} \frac{\lambda'_\beta(u)}{\lambda_\beta(u)} - \int_W \lambda'_\beta(u) \mathrm{d}u$$

optimal for Poisson (likelihood).

Which f makes

$$e_f(\beta) = \sum_{u \in \mathbf{X}} f(u) - \int_W f(u) \lambda_\beta(u) \mathrm{d}u$$

optimal for Cox point process (positive dependence between points) ?

Optimal first-order estimating equation

Optimal choice of f: smallest variance

$$\mathbb{V}\mathrm{ar}\hat{\beta} = V_f = S_f^{-1}\Sigma_f S_f^{-1}$$

where

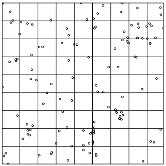
$$S_f = -\mathbb{E}\frac{\mathrm{d}}{\mathrm{d}\beta^{\mathsf{T}}}e_f(\beta) \quad \Sigma_f = \mathbb{V}\mathrm{ar}e_f(\beta)$$

Possible to obtain optimal f as solution of certain Fredholm integral equation.

Numerical solution of integral equation leads to estimating function of quasi-likelihood type.

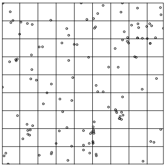
Quasi-likelihood

Approximate solution of Fredholm integral equation using numerical quadrature: Riemann sum dividing W into cells C_i with representative points u_i.



Quasi-likelihood

Approximate solution of Fredholm integral equation using numerical quadrature: Riemann sum dividing W into cells C_i with representative points u_i.



Resulting estimating function is *quasi-likelihood*

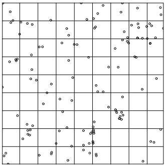
$$(N - \mu)V^{-1}D$$

based on

$$N = (N_1, \ldots, N_m), \quad N_i = 1[\mathbf{X} \text{ has point in } C_i].$$

Quasi-likelihood

Approximate solution of Fredholm integral equation using numerical quadrature: Riemann sum dividing W into cells C_i with representative points u_i.



Resulting estimating function is *quasi-likelihood*

$$(N - \mu)V^{-1}D$$

based on

$$N = (N_1, \ldots, N_m), \quad N_i = 1[\mathbf{X} \text{ has point in } C_i].$$

μ mean of N:

$$\mu_i = \mathbb{E}N_i = \lambda_\beta(u_i)|C_i| \text{ and } D = \left[\mathrm{d}\mu(u_i)/\mathrm{d}\beta_l\right]_{il}$$

V covariance of N (involves covariance of random intensity):

$$V_{ij} = \mathbb{C}\mathrm{ov}[N_i, N_j] = \mu_i 1[i = j] + \mu_i \mu_j \mathbb{C}\mathrm{ov}[\Lambda(u_i), \Lambda(u_j)]$$

Results with composite likelihood and quasi-likelihood

species		$\widehat{\beta}$
Loncocharpus	CL	$-6.49 - 0.021\texttt{Nmin} - 0.11\texttt{P} - 0.59\texttt{pH} - 0.11\texttt{twi}$ $(81.06^*, 7.45^*, 58.78, 282.89^*, 53.19^*) \times 10^{-3}$
	QL	$-6.49 - 0.023\texttt{Nmin} - 0.12\texttt{P} - 0.55\texttt{pH} - 0.084\texttt{twi}$ $(80.15^*, 6.95^*, 55.23^*, 266.10^*, 45.47) \times 10^{-3}$
Capparis	CL	$-5.07 + 0.028\texttt{ele} - 1.10\texttt{grad} + 0.0043\texttt{K}$ $(79.54^*, 9.98^*, 1200.36, 1.16^*) \times 10^{-3}$
	QL	$-5.10 + 0.019\texttt{ele} - 2.50\texttt{grad} + 0.0039\texttt{K}$ $(77.77^*, 8.86^*, 935.02^*, 1.02^*) \times 10^{-3}$

Estimated standard errors always smallest for QL. Covariate `grad` significant according to QL but not for CL.

References

Waagepetersen (2007). An estimating function approach to inference for inhomogeneous Neyman-Scott processes, *Biometrics*.

Waagepetersen, R. (2007). Estimating functions for inhomogeneous spatial point processes with incomplete covariate data, *Biometrika*.

Jalilian, Guan and Waagepetersen (2012). Decomposition of variance for spatial Cox processes, *Scandinavian Journal of Statistics*, to appear.

Guan, Jalilian and Waagepetersen (2012). Optimal first order estimating functions for spatial point processes, submitted.

Baddeley, Couerjolly, Rubak and Waagepetersen (2012). A logistic regression estimating function for spatial Gibbs point processes, in preparation.

Thanks for your attention !