Indian Buffet Epidemics

Ashley Ford

RSC 2010

Outline

1. Epidemic Models and Inference
   - SIR Models
2. Indian Buffet Epidemics
   - MCMC Inference

SIR Models

$$S \xrightarrow{\beta} I \xrightarrow{\gamma} R$$

$$N_t^S \qquad N_t^I \qquad N_t^R$$

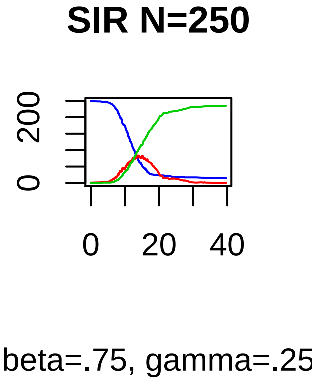
$$N = N_t^S + N_t^I + N_t^R$$

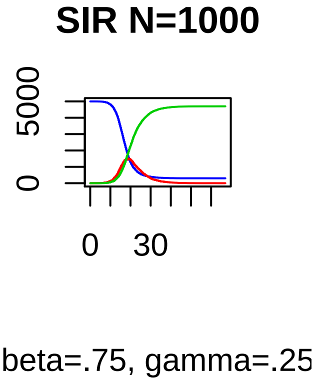
$$P(N_{t+dt}^S = j - 1 | N_t^S = j) = \beta N_t^S N_t^I dt$$

$$P(N_{t+dt}^I = j + 1 | N_t^I = j) = \beta N_t^S N_t^I dt$$

$$P(N_{t+dt}^I = j - 1 | N_t^I = j) = \gamma N_t^I dt$$

SIR examples





Inference

- What data is available ?
  - Epidemic complete ?
  - Infection times ?
- MLE well known with full data
  - see Andersson and Britton (2000)
- Martingale estimator Becker and Hasofer (1997)
- MCMC estimates O'Neill and Roberts (1999)

Inference

- What data is available ?
  - Epidemic complete ?
  - Infection times ?
- MLE well known with full data
  - see Andersson and Britton (2000)
- Martingale estimator Becker and Hasofer (1997)
- MCMC estimates O'Neill and Roberts (1999)

Inference

- What data is available ?
  - Epidemic complete ?
  - Infection times ?
- MLE well known with full data
  - see Andersson and Britton (2000)
- Martingale estimator Becker and Hasofer (1997)
- MCMC estimates O'Neill and Roberts (1999)

Inference

- What data is available ?
  - Epidemic complete ?
  - Infection times ?
- MLE well known with full data
  - see Andersson and Britton (2000)
- Martingale estimator Becker and Hasofer (1997)
- MCMC estimates O'Neill and Roberts (1999)

Indian Buffet Epidemics

- Need a model between homogeneous mixing and over complex models.
- Aim to fit the heterogeneity with two or three parameters that measure the departure from homogeneity.

Places and People

- Model heterogeneity in an epidemic amongst N people
- Each person belongs to 1 or more of many classes
  - e.g. households, schools, clubs, buses etcetera
- The classes are not specified
- A prior is put on class membership
  - represented as an $N \times K$ binary matrix Z
- An Indian Buffet Process

Places and People

- Model heterogeneity in an epidemic amongst N people
- Each person belongs to 1 or more of many classes
  - e.g. households, schools, clubs, buses etcetera
- The classes are not specified
- A prior is put on class membership
  - represented as an $N \times K$ binary matrix Z
- An Indian Buffet Process

Places and People

- Model heterogeneity in an epidemic amongst N people
- Each person belongs to 1 or more of many classes
  - e.g. households, schools, clubs, buses etcetera
- The classes are not specified
- A prior is put on class membership
  - represented as an $N \times K$ binary matrix Z
- An Indian Buffet Process

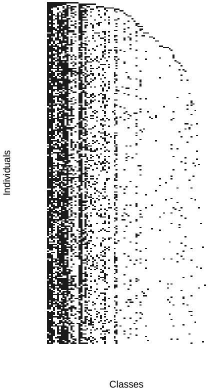
Indian Buffet Process

- Introduced by Griffiths and Ghahramani (2005)
- For each k $\psi_k$is the probability that an individual is in class k
- $\psi_k \sim Beta(\alpha/K, 1)$with α being the strength parameter of the IBP.
- The model for Z is: $z_{ik}|\psi_k \sim Bernoulli(\psi_k)$ independently
- The process is obtained as $K \rightarrow \infty$

A culinary metaphor

- N customers enter a restaurant one after another.
- The jth customer selects each dish with probability m_k/j
  - where m_k is the number of previous customers who have chosen a dish.
- He then tries $\text{Poisson}(\alpha/j)$ new dishes.

Indian Buffet Process example



IBP Z generated with $N = 260$, $K = 260$, $\alpha = 15$

Indian Buffet Epidemic

- The state of individual j is at time t is $x_{j,t} \in$ {S,I,R} .
- Given Z , infections are independent with transition rates given by
  - $P(x_{j,t+dt} = I \mid x_{j,t} = S, Z) = \sum z_{jk} \lambda_k N^I_{k,t}\, dt$
  - where $N^I_{k,t}$is the number that are in classs k and infective at time t and $\lambda_k$is the infection rate within group k.
- $N^I_{k,t} = \sum_j z_{jk} \mathbf{1}(x_{j,t} = I)$
- A basic model has λ_k the same for all k.
- Intuitively it is reasonable to assume a greater per person infection rate in a small group such as a household
  - so take $\lambda_k = \lambda N_k^{-\nu}$

Indian Buffet Epidemic

- The state of individual j is at time t is $x_{j,t} \in$ {S,I,R} .
- Given Z , infections are independent with transition rates given by
    - $P(x_{j,t+dt} = I \mid x_{j,t} = S, Z) = \sum z_{jk} \lambda_k N^I_{k,t}\, dt$
    - where $N^I_{k,t}$is the number that are in classs k and infective at time t and $\lambda_k$is the infection rate within group k.
- $N^I_{k,t} = \sum_j z_{jk} \mathbf{1}(x_{j,t} = I)$
- A basic model has λ_k the same for all k.
- Intuitively it is reasonable to assume a greater per person infection rate in a small group such as a household
    - so take $\lambda_k = \lambda N_k^{-\nu}$

MCMC Inference

- Augmented data
- Parameterisation
- Proposal

Summary

- The limitations of existing epidemic models have been explored.
- A new model for epidemics incorporating heterogenity has been introduced.
- Initial steps towards inference taken.

- Planned developments
  - Develop MCMC algorithms
  - Apply to real data
- Questions