SPATIAL DIFFUSION ANALYSIS

Example Application Areas

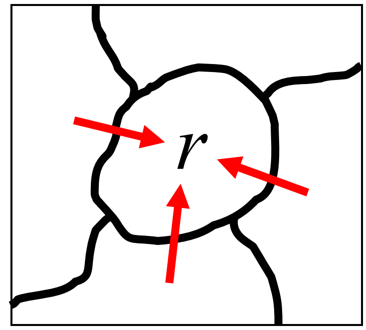
- Diffusion of Information
- Diffusion of Toxic Wastes
- Spread of Infectious Diseases

Product Adoption Example

(http://www.seas.upenn.edu~tesmith)

- Basic Model
- Steady State Analysis
- Parameter Estimation
- Philadelphia Application

BASIC MODEL



Regions: $r \in \mathbf{R} = \{r_1, .., r_R\}$

Adoptions: $(r_n, n = 0, 1, .., N)$

Adoption Frequencies: $f_n = \left[f_n(r) : r \in \mathbf{R} \right]$

- **Mixture Distribution**

$$p_n(r \mid r_0, r_1, .., r_{n-1}) = \lambda p_c(r \mid f_n) + (1-\lambda) p_0(r)$$

- **Contact Model**

$$p_c(r \mid f_n) = \sum_{s \in \mathbf{R}} \frac{M_r \exp(-\theta c_{sr})}{\sum_{v \in \mathbf{R}} M_v \exp(-\theta c_{sv})} f_n(s)$$

- **Intrinsic Model**

$$p_0(r) = \frac{M_r \exp(-\beta x_r)}{\sum_{s \in \mathbf{R}} M_s \exp(-\beta x_s)}$$

STEADY STATE ANALYSIS

- **State Probability Mapping**

$$p(f) = \lambda P_c f + (1-\lambda) p_0$$

- **Fixed Point Property**

$$f = p(f) = \lambda P_c f + (1-\lambda) p_0$$

$$\Rightarrow \quad f^* = (1-\lambda)(I - \lambda P_c)^{-1} p_0$$

- **Convergence to Steady State**

$$\Pr\left(\lim_{n\to\infty} f_n = f^*\right) = 1$$

- **Rate of Convergence**

$$|f_n - f^*| = O\left(\exp^{(\lambda-1)t_n}\right) \qquad t_n = \sum_{m=0}^{n} \left(\tfrac{1}{m}\right)$$

MAXIMUM LIKELIHOOD

- **Observed Data:** $y = (y_0, y_1, .., y_N)$
- **Log Likelihood Function**

$$L(\beta, \lambda, \theta \mid y) = \log p_0(y_0) + \sum_{n=1}^{N} \log p_n(y_n \mid f_n)$$

where:

$$\log p_n(y_n \mid f_n) = \log\left[\lambda p_\theta(y_n \mid f_n) + (1-\lambda) p_\beta(y_n)\right]$$

- **Problem:** Can have

$$p_\theta(y_n \mid f_n) < p_\beta(y_n) \; , \; n = 1, .., N$$

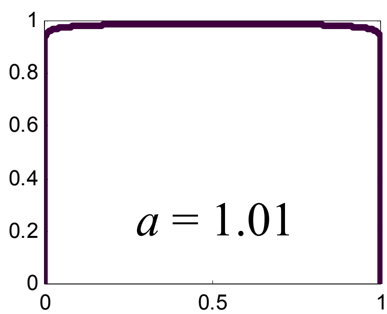
- **Example:** $R = 18, \; J = 2, \; N = 200, \; c_{rs} = d_{rs}$

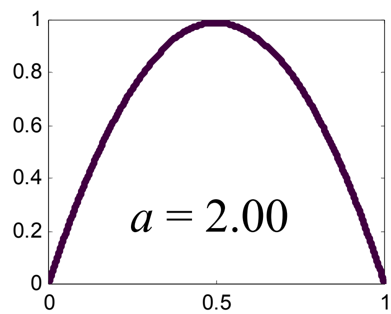
Param	Value	Estimate	P-value
β_1	1.0	1.00038	0.00002
β_2	-2.0	-2.17340	0.00000
λ	0.30	-1805646	0.00000
θ	10.0	-2838.631	0.99999

BAYESIAN ESTIMATION

- **Prior Distributions:**

$$\pi(\lambda) \propto \lambda^{a-1}(1-\lambda)^{a-1} \qquad \pi(\beta), \pi(\theta) \propto 1$$





- **Maximum Aposteriori (MAP) Estimates**

$$\Phi(\beta,\lambda,\theta \mid y) = L(\beta,\lambda,\theta \mid y) + (a-1)\left[\log\lambda + \log(1-\lambda)\right]$$

Param	Value	Estimate	P-value
β_1	1.0	0.99939	0.00002
β_2	-2.0	-2.17165	0.00000
λ	0.30	0.00001	0.92034
θ	10.0	153.963	0.99999

FULL BAYES MODEL

- **Conditional Probability Model:**

$$p(y \mid \lambda, \beta, \theta) = p(y_0 \mid \beta) \prod_{n=1}^{N} p(y_n \mid f_n, \lambda, \beta, \theta)$$

- **Prior Distributions:**

$$\beta \sim N\left(0, vI\right) \;\Rightarrow\; \pi(\beta) \propto e^{-v\beta'\beta/2}$$

$$\theta \sim \Gamma(b, c) \;\Rightarrow\; \pi(\theta) \propto \theta^{b-1} e^{-c\theta}$$

- **Posterior Distributions:**

$$p(\lambda, \beta, \theta \mid y) \propto p(y \mid \lambda, \beta, \theta)\pi(\lambda)\pi(\beta)\pi(\theta)$$

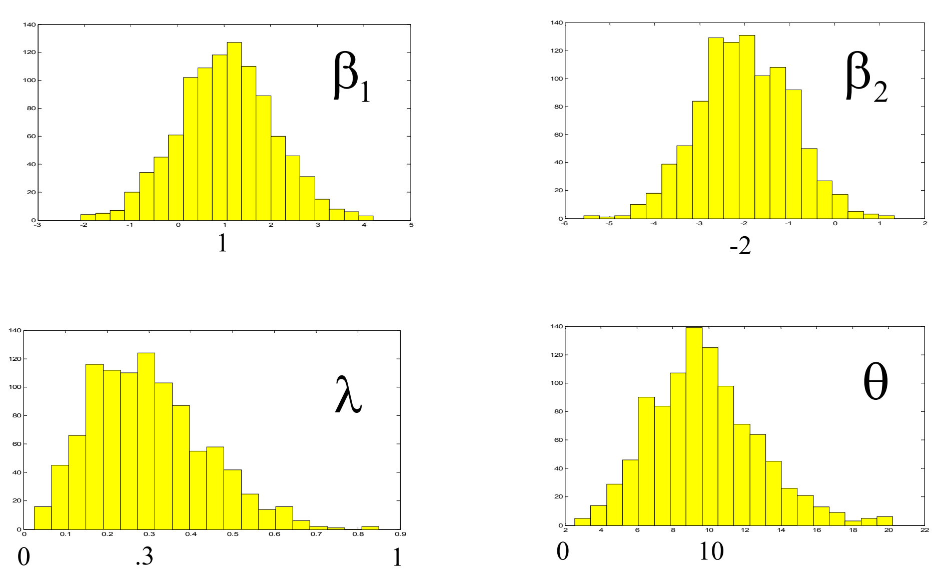
$$\Rightarrow\; p(\lambda \mid \beta_0, \theta_0, y) \propto p(\lambda, \beta_0, \theta_0 \mid y)$$

$$p(\beta \mid \lambda_0, \theta_0, y) \propto p(\lambda_0, \beta, \theta_0 \mid y)$$

$$p(\theta \mid \lambda_0, \beta_0, y) \propto p(\lambda_0, \beta_0, \theta \mid y)$$

BAYES MONTE CARLO

- **Gibbs Sampling Procedure:**
  - **Start with any initial values** $(\lambda_0, \beta_0, \theta_0)$
  - **Sample new** $\lambda_1 \sim p(\lambda \mid \beta_0, \theta_0, y)$
  - **Sample new** $\beta_1 \sim p(\beta \mid \lambda_1, \theta_0, y)$
  - **Sample new** $\theta_1 \sim p(\theta \mid \lambda_1, \beta_1, y)$
  - **Now start with** $(\lambda_1, \beta_1, \theta_1)$ **and continue**
- **Save final values** $\left[(\lambda_m, \beta_m, \theta_m) : m = M_0, .., M_1\right]$
- **Plot marginal sampling distributions**



SIMULATION RESULTS

BETA 1

Size	Mean	Medn	Stdev	% < 0
100	1.173	1.057	0.615	0
200	1.102	1.029	0.413	0
500	1.077	1.017	0.372	0
1000	1.012	1.002	0.159	0
2000	1.008	0.999	0.098	0

LAMBDA

Size	Mean	Medn	Stdev	% < .01
100	0.264	0.266	0.130	0.029
200	0.259	0.261	0.106	0.005
500	0.269	0.263	0.096	0.001
1000	0.274	0.274	0.071	0
2000	0.280	0.275	0.061	0

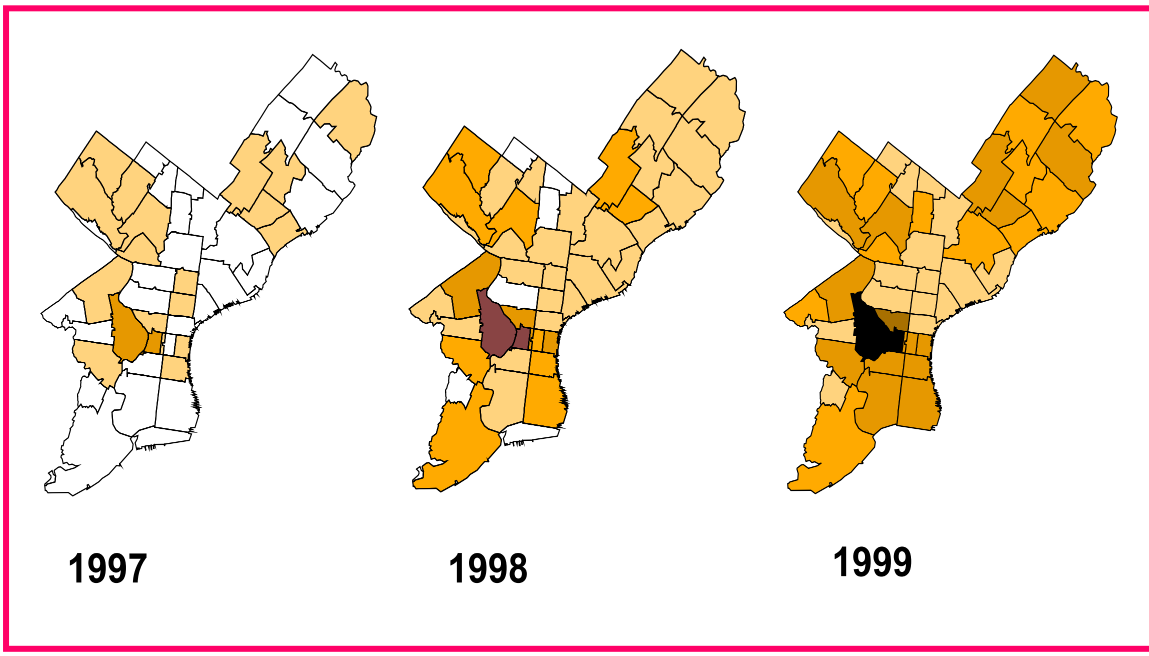
THETA

Size	Mean	Medn	Stdev	% < 0
100	16.38	9.67	169.2	0.041
200	11.12	9.13	123.7	0.020
500	5.11	9.07	130.4	0.002
1000	9.31	9.42	2.21	0
2000	9.40	9.48	1.78	0

PHILADELPHIA APPLICATION

- **First purchases at Netgrocer.com**

(*N* = 1288 over 3 yrs., *R* = 46 zipcode areas)



➡ **Concentrated in University Area**

PHILADELPHIA DATA

- **INTRINSIC VARIABLES**

Variable	Description
BDR5	% of Housing units with more than 5 Bedrooms
COLDEG	% of over 25 year-olds with College Degrees
DIWK	% of Households with both parents working.
ELDERLY	% of population over 65 years old.
FAMLARG	% of Households with more that 5 members
SOLO	% of Households with exactly one member
SUPMAS	Number of Supermarkets per person

- **CONTACT COSTS = Centroid Distances**

ESTIMATION RESULTS

Variable	Estimate	P-Value
BDR5	6.045	0.1302
COLDEG	-3.952	0.1637
DIWK	-0.467	0.5611
ELDERLY	-7.728	0.0182
FAMLARG	-13.58	0.0041
SOLO	7.781	0.0094
SUPMAS	-158.74	0.6186
LAMBDA	0.678	< .0000
THETA	1195.9	0.9996

- **Lambda+Theta shows strong local contacts**

  ($\theta \rightarrow -\infty \Rightarrow$ P-value not meaningful)

- **All significant values are consistent with the student populations where adoptions are concentrated.**