Game-changers:
Detecting shifts in the flow of campaign contributions

Matthew Blackwell
University of Rochester

March 8th, 2013
APWG
Why not polls?

1. Cheap talk
2. Data (un)availability
Why not polls?

1. Lack of variation
Why not polls?

1. Lack of variation

2. Cheap talk
Why not polls?

1. Lack of variation

2. Cheap talk

3. Data (un)availability
NAME OF COMMITTEE (in Full)  
Obama for America

A. Full Name (Last, First, Middle Initial)  
Sharon Anderson  
Mailing Address 1668 finwick dr  
City piafstown  
State NC  
Zip Code 27040-9043  
FEC ID number of contributing federal political committee.  
C  
Name of Employer  
The Norman Group  
Occupation  
Intl Consultant  
Receipt For: 2012  
Primary  
Other (specify) ▼  
Election Cycle-to-Date ▼  
Amount of Each Receipt this Period  
11.00  
Transaction ID : C19176830  
Date of Receipt  
08/12/2012

B. Full Name (Last, First, Middle Initial)  
Riaz Hussain  
Mailing Address 540 N Webster Ave  
City Scranton  
State PA  
Zip Code 18510  
FEC ID number of contributing federal political committee.  
C  
Name of Employer  
University of Scranton  
Occupation  
Professor  
Receipt For: 2012  
Primary  
Other (specify) ▼  
Election Cycle-to-Date ▼  
Amount of Each Receipt this Period  
35.00  
Transaction ID : C20196560  
Date of Receipt  
08/30/2012

C. Full Name (Last, First, Middle Initial)  
Dave Baird  
Mailing Address 1376 Lincoln St  
Transaction ID : C20090710  
Date of Receipt  
08/30/2012
A measurement question

When do campaigns take off or fall flat?
A measurement question

When do campaigns take off or fall flat?

When do campaign contributions take off or fall flat?
When do campaigns take off or fall flat?

When do campaign contributions take off or fall flat?

Tools: Bayesian nonparametric model for overdispersed count data.
Why contributions?
Why contributions?

1. Lots of variation
Why contributions?

1. Lots of variation

2. Costly participation
Why contributions?

1. Lots of variation

2. Costly participation

3. Data availability
Why changepoint models?
The challenges

Modeling daily contribution counts
The challenges

Modeling daily contribution counts

Choosing the number of changepoints
Overdispersion in campaign contributions
Bayesian model for overdispersed counts

For observations $t$ in $\{1, \ldots, T\}$:

$$[y_t|\eta_t, \beta, \rho, X] \sim \text{Poisson}(\eta_t \lambda_t)$$ (data)
Bayesian model for overdispersed counts

For observations $t$ in $\{1, \ldots, T\}$:

$$[y_t|\eta_t, \beta, \rho, X] \sim \text{Poisson}(\eta_t \lambda_t)$$  \hspace{1cm} \text{(data)}

$$\lambda_t = \exp(X_t \beta)$$  \hspace{1cm} \text{(link function)}
For observations $t$ in $\{1, \ldots, T\}$:

$$[y_t|\eta_t, \beta, \rho, X] \sim \text{Poisson}(\eta_t \lambda_t)$$ (data)

$$\lambda_t = \exp(X_t \beta)$$ (link function)

$$[\eta_t|\rho] \sim \text{Gamma}(\rho, \rho)$$ (random effect)
Bayesian model for overdispersed counts

For observations $t$ in $\{1, \ldots, T\}$:

$$[y_t|\eta_t, \beta, \rho, X] \sim \text{Poisson}(\eta_t \lambda_t)$$ (data)

$$\lambda_t = \exp(X_t \beta)$$ (link function)

$$[\eta_t|\rho] \sim \text{Gamma}(\rho, \rho)$$ (random effect)

marginal distribution of $y$:

$$[y_t|\beta, \rho, X] \sim \text{NegBin}(\rho, \rho/(\rho + \lambda_t))$$
Generalize to a mixture model

$[y_t | s_t, \eta_t, \beta, \rho, X] \sim \text{Poisson}(\eta_t \lambda_t)$  (data)
Generalize to a mixture model

\[ y_t | s_t, \eta_t, \beta, \rho, X ] \sim \text{Poisson}(\eta_t \lambda_t) \]  
(data)

regimes \quad s_t = k 
(1, \ldots, K)
Generalize to a mixture model

\[ y_t | s_t, \eta_t, \beta, \rho, X \sim \text{Poisson}(\eta_t \lambda_t) \]  

\( s_t = k \)  

\( \lambda_t = \exp(X_t \beta_k) \)  

\( \eta_t | \rho, s_t \sim \text{Gamma}(\rho s_t, \rho s_t) \)  

Pr(\( s_{t+1} = k \)) = \( p_k \)  

Pr(\( s_{t+1} = k + 1 \)) = 1 - \( p_k \)  

Pr(\( s_{t+1} = j \)) = 0  

(\( \forall j \notin \{k, k+1\} \))
Generalize to a mixture model

\[ [y_t|s_t, \eta_t, \beta, \rho, X] \sim \text{Poisson}(\eta_t \lambda_t) \]  
(data)

\[ \lambda_t = \exp(X_t \beta_k) \]  
(link function)

\[ [\eta_t|\rho, s_t] \sim \text{Gamma}(\rho_k, \rho_k) \]  
(random effect)

regimes
\( s_t = k \)

(1, ..., K)

Pr(\( s_{t+1} = k \) | \( s_t = k \)) = \( p_k \)

Pr(\( s_{t+1} = k + 1 \) | \( s_t = k \)) = 1 - \( p_k \)

Pr(\( s_{t+1} = j \) | \( s_t = k \)) = 0 \quad (\forall j \notin \{k, k+1\})
Generalize to a mixture model

\[ [y_t | s_t, \eta_t, \beta, \rho, X] \sim \text{Poisson}(\eta_t \lambda_t) \] (data)

\[ \lambda_t = \exp(X_t \beta_k) \] (link function)

\[ [\eta_t | \rho, s_t] \sim \text{Gamma}(\rho_k, \rho_k) \] (random effect)

\[ Pr(s_{t+1} = k | s_t = k) = p_k \]
Generalize to a mixture model

\[ [y_t|s_t, \eta_t, \beta, \rho, X] \sim \text{Poisson}(\eta_t \lambda_t) \]

(data)

\[ \lambda_t = \exp(X_t \beta_k) \]

(link function)

\[ [\eta_t|\rho, s_t] \sim \text{Gamma}(\rho_k, \rho_k) \]

(random effect)

\[ \Pr(s_{t+1} = k|s_t = k) = p_k \]
\[ \Pr(s_{t+1} = k + 1|s_t = k) = 1 - p_k \]
Generalize to a mixture model

\[
[y_t|s_t, \eta_t, \beta, \rho, X] \sim \text{Poisson}(\eta_t \lambda_t) \quad \text{(data)}
\]

\[
\lambda_t = \exp(X_t \beta_k) \quad \text{(link function)}
\]

\[
[\eta_t|\rho, s_t] \sim \text{Gamma}(\rho_k, \rho_k) \quad \text{(random effect)}
\]

\[
\begin{align*}
\Pr(s_{t+1} = k | s_t = k) &= p_k \\
\Pr(s_{t+1} = k + 1 | s_t = k) &= 1 - p_k \\
\Pr(s_{t+1} = j | s_t = k) &= 0 \\
\end{align*}
\quad (\forall j \notin \{k, k + 1\})
\]
Traditional changepoint models

Regimes

1  2  3  4

Units

$1 - p_1$, $\ldots$, $N$

Must be in the last regime
Traditional changepoint models

Regimes

1. $\beta_1, \rho_1$
2. $\beta_2, \rho_2$
3. $\beta_3, \rho_3$
4. $\beta_4, \rho_4$

Each regime must be in the last regime.
Traditional changepoint models

Regimes:
1. $(\beta_1, \rho_1)$
2. $(\beta_2, \rho_2)$
3. $(\beta_3, \rho_3)$
4. $(\beta_4, \rho_4)$

Units:
1. $p_1$ - $1 - p_1$
Traditional changepoint models

Regimes

$(\beta_1, \rho_1)$  $(\beta_2, \rho_2)$  $(\beta_3, \rho_3)$  $(\beta_4, \rho_4)$

Units

$\cdot$
Traditional changepoint models

Regimes

- $(\beta_1, \rho_1)$
- $(\beta_2, \rho_2)$
- $(\beta_3, \rho_3)$
- $(\beta_4, \rho_4)$

Units

1  2

Must be in the last regime
Traditional changepoint models

Regimes

$(\beta_1, \rho_1)$

$(\beta_2, \rho_2)$

$(\beta_3, \rho_3)$

$(\beta_4, \rho_4)$

Units

1

2

3

$p_1$

$1 - p_1$
Traditional changepoint models

Regimes

$(\beta_1, \rho_1)$

$(\beta_2, \rho_2)$

$(\beta_3, \rho_3)$

$(\beta_4, \rho_4)$

Units

1

2

3

changepoint

$\mathbf{p} = 1 - \mathbf{p}$

Must be in the last regime
Traditional changepoint models

Regimes

\((\beta_1, \rho_1)\)

\((\beta_2, \rho_2)\)

\((\beta_3, \rho_3)\)

\((\beta_4, \rho_4)\)

Units

1

2

3

⋯

changepoint

Must be in the last regime
Traditional changepoint models

Regimes

\((\beta_1, \rho_1)\) \n\((\beta_2, \rho_2)\) \n\((\beta_3, \rho_3)\) \n\((\beta_4, \rho_4)\)

Units

1 \n2 \n3 \n...
N

changepoint

\(p_1 - \rho_1\)
Traditional changepoint models

Regimes

- $(\beta_1, \rho_1)$
- $(\beta_2, \rho_2)$
- $(\beta_3, \rho_3)$
- $(\beta_4, \rho_4)$

Units

- 1
- 2
- 3
- \(\ldots\)
- N

须在最后一个阶段

changepoint
Bayesian nonparametric priors

- Model assumptions: $y_t \sim G$ i.i.d. from an unknown distribution $G$. 
Bayesian nonparametric priors

- Model assumptions: $y_t \sim G$ i.i.d. from an unknown distribution $G$.

- Parametric structure on our priors puts limitations on the posterior inferences.
Bayesian nonparametric priors

- Model assumptions: $y_t \sim G$ i.i.d. from an unknown distribution $G$.

- Parametric structure on our priors puts limitations on the posterior inferences.

- Bayesian nonparametrics: priors over distributions and, thus, an infinite number of parameters.
Dirichlet process prior

- clusters units into a countably infinite set of groups.
- Obviously we only observe a finite number of these groups, but the number is determined by the data and the prior, not exclusively from the prior.
Dirichlet process prior

- Dirichlet process prior clusters units into a countably infinite set of groups.
• Dirichlet process prior clusters units into a countably infinite set of groups.

• Obviously we only observe a finite number of these groups, but the number is determined by the data and the prior, not exclusively from the prior.
Dirichlet process prior clusters units into a countably infinite set of groups.

Obviously we only observe a finite number of these groups, but the number is determined by the data and the prior, not exclusively from the prior.

\[
\begin{align*}
\Pr(s_{t+1} = k \mid s_t = k) &= \frac{n_k}{t-1+b} \\
\Pr(s_{t+1} = k + 1 \mid s_t = k) &= \frac{b}{t-1+b}
\end{align*}
\]
Dirichlet process prior

Regimes (infinite)

\((\beta_1, \rho_1)\)

\((\beta_2, \rho_2)\)

\((\beta_3, \rho_3)\)

\(\cdots\)

\(\cdots\)

Units

\(n_1/(t - 1 + b)\)

\(b/(t - 1 + b)\)

changepoint

\(n_1/\cdots/n_N(\beta_1\cdots, \rho_1\cdots)\)

Number of regimes can vary
Dirichlet process prior

Regimes
1
2
3
(...)

(infinite)

Units

\( \frac{n - \beta}{t - \beta + b} \)

\( \frac{b}{t - \beta + b} \)

Number of regimes can vary
Dirichlet process prior

Regimes (infinite)

$(\beta_1, \rho_1)$ $(\beta_2, \rho_2)$ $(\beta_3, \rho_3)$ $\cdots$ $\cdots$ $\cdots$

Units

$n/(t - \tau + b)$ $b/(t - \tau + b)$

Number of regimes can vary
Dirichlet process prior

Regimes

(\(\beta_1, \rho_1\))  (\(\beta_2, \rho_2\))  (\(\beta_3, \rho_3\))  \(\ldots\)  \(\ldots\)  \(\ldots\)

(infinite)

Units

(\(n/(t - \beta + b)\))  (\(b/(t - \beta + b)\)
Dirichlet process prior

Regimes (infinite)

Units

1

\((\beta_1, \rho_1)\) \hspace{1cm} \((\beta_2, \rho_2)\) \hspace{1cm} \((\beta_3, \rho_3)\) \hspace{1cm} \ldots \hfill \ldots \hfill \ldots \hfill \ldots
Dirichlet process prior

Regimes (infinite)

Units

(\(\beta_1, \rho_1\))  (\(\beta_2, \rho_2\))  (\(\beta_3, \rho_3\)) ...

Number of regimes can vary

\(n/(t - \delta + b)\)  \(b/(t - \delta + b)\)
Dirichlet process prior

Regimes

\((\beta_1, \rho_1)\) \(\quad (\beta_2, \rho_2)\) \(\quad (\beta_3, \rho_3)\) \(\quad \cdots \quad \cdots \quad \cdots \)

(\text{infinite})

Units

1 \quad 2 \quad 3

\(n_1/(t - 1 + b)\) \(b/(t - 1 + b)\)

Number of regimes can vary
Dirichlet process prior

Regimes (infinite)

$((\beta_1, \rho_1), (\beta_2, \rho_2), (\beta_3, \rho_3), \ldots)$

Units $1, 2, 3, \ldots$

changepoint
Dirichlet process prior

Regimes (infinite)

\((\beta_1, \rho_1)\) \(\rightarrow\) \((\beta_2, \rho_2)\) \(\rightarrow\) \((\beta_3, \rho_3)\) \(\rightarrow\) \(\cdots\)

Units

\(1\) \(\rightarrow\) \(2\) \(\rightarrow\) \(3\) \(\rightarrow\) \(\cdots\)

changepoint

Number of regimes can vary
Dirichlet process prior

Regimes
(infinite)

Units

(changepoint)

\((\beta_1, \rho_1)\) \(\to\) \((\beta_2, \rho_2)\) \(\to\) \((\beta_3, \rho_3)\) \(\to\) \(\cdots\) \(\cdots\) \(\cdots\)

\(\cdots\)
Dirichlet process prior

Regimes (infinite)

Units

Number of regimes can vary

$(\beta_1, \rho_1)$

$(\beta_2, \rho_2)$

$(\beta_3, \rho_3)$

$\cdots$

$(\beta_{10}, \rho_{10})$

1

2

3

$\cdots$

$N$

changepoint

Units

Number of regimes can vary
Your lunch is never free

- DPP has a rich-get-richer property:

\[
\Pr(s_{t+1} = k \mid s_t = k) = \frac{n_k}{t - 1 + b}
\]
DPP has a rich-get-richer property:

$$\Pr(s_{t+1} = k \mid s_t = k) = \frac{n_k}{t - 1 + b}$$

No free lunch theorem: All nonparametric priors place assumptions on the clustering algorithm and no algorithm is optimal across the space of all problems.
Monte Carlo evidence
Monte Carlo evidence

Negative Binomial

Poisson

Probability of Changepoint

0.0 0.2 0.4 0.6 0.8 1.0

0 50 100 150 200

0.0 0.2 0.4 0.6 0.8 1.0

0 50 100 150 200

0 50 100 150 200

0.0 0.2 0.4 0.6 0.8 1.0

0 50 100 150 200
The rise and fall of Herman Cain

Probability of Changepoint

Jan 11

Apr 11

Jul 11

Oct 11

Jan 12

0.0

0.2

0.4

0.6

0.8

Number of Contributions

Jan 11

Apr 11

Jul 11

Oct 11

Jan 12

0

100

200

300

400

500

Fox News Debate

Announces candidacy

Wins FL straw poll

Misconduct allegations

Suspends candidacy
The rise and fall of Herman Cain

Probability of Changepoint

Number of Contributions

Jan 11
Apr 11
Jul 11
Oct 11
Jan 12

0
100
200
300
400
500

0.0
0.2
0.4
0.6
0.8

Fox News Debate
Announces candidacy
Wins FL straw poll
Misconduct allegations
Suspends candidacy
The rise and fall of Herman Cain
The rise and fall of Herman Cain

Probabilty of Changepoint

Number of Contributions

Jan 11  Apr 11  Jul 11  Oct 11  Jan 12

0  100  200  300  400  500
The rise and fall of Herman Cain
The rise and fall of Herman Cain
The rise and fall of Herman Cain

Probability of Changepoint

Number of Contributions

Fox News Debate
Announces candidacy
Wins FL straw poll

Jan 11
Apr 11
Jul 11
Oct 11
Jan 12

0
100
200
300
400
500
0.0
0.2
0.4
0.6
0.8
The rise and fall of Herman Cain

Probability of Changepoint

Jan 11

Apr 11

Jul 11

Oct 11

Jan 12

0.0

0.2

0.4

0.6

0.8

Number of Contributions

Jan 11

Apr 11

Jul 11

Oct 11

Jan 12

0

100

200

300

400

500

Fox News Debate

Announces candidacy

Wins FL straw poll

Misconduct allegations

Suspends candidacy
More attention around changepoints
The path forward.
The path forward.

- Run on all (digitized) Congressional races to find more systematic variation.
The path forward.

1. Run on all (digitized) Congressional races to find more systematic variation.

2. Compare changepoints for time-series of different types of voters, PACs.
The path forward.

1. Run on all (digitized) Congressional races to find more systematic variation.

2. Compare changepoints for time-series of different types of voters, PACs.

3. Generalize the Bayesian nonparametric approach beyond count data.