Persuasion Profiling and Streaming Data Analysis and . . .

Dr. M.C.Kaptein

October 22, 2015

Section 1

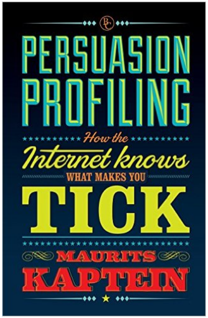
Background

Education

- MSc. Economic Psychology, Tilburg University
- PdEng. User System Interaction, University of Eindhoven
- Ph.D. Industrial Design, University of Eindhoven & Stanford University
- Post Doc. Marketing, Aalto School of Economics, Helsinki

Current appointments

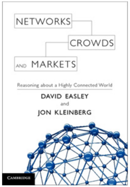
- Assistant Professor, Artificial Intelligence, Radboud University Nijmegen
- Founder & Chief Scientist, PersuasionAPI & Science Rockstars, Amsterdam / Barneveld. Acquired by Webpower b.v..
- Speaker for The Next Speaker
- Author: “Persuasion Profiling” (in Dutch “Digitale Verleiding”)



Section 2

Persuasion

Persuasion in E-Commerce

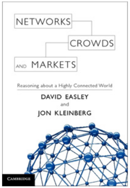


Networks, Crowds & Markets

Easley & Kleinberg

Recommended price: $ 14.99

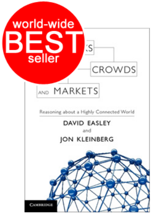
Persuasion in E-Commerce



Networks, Crowds & Markets

Easley & Kleinberg

Recommended price: $ 14.99



Networks, Crowds & Markets

Easley & Kleinberg

Recommended price: $ 14.99

Main Effects of Persuasion

- Average effect of the little "button": Willingness to pay increase by $> 30\%$
- Similar effects in different studies: Probability of purchase increased by 5 to 25%

Distinct Persuasion Strategies

- Scarcity
- Authority
- Social proof
- Liking
- Reciprocity
- Commitment
- ...

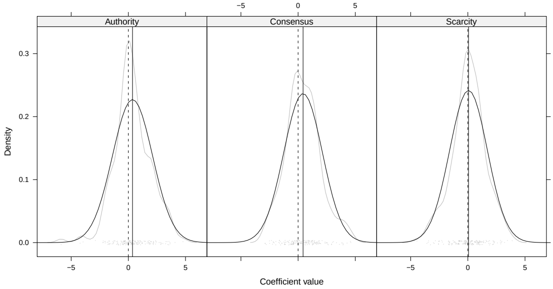
Estimating individual level effects of persuasion[1]

$$y_{jbq} \sim \mathcal{N}(X_{jb}\beta_j + \alpha_b + \eta_q, \sigma^2_{err}) \tag{1}$$

with
$\beta_j \sim \mathcal{N}(\bar{\beta}, \Sigma_\beta)$ for $j = 1, ..., J = 179$ subjects
$\alpha_b \sim \mathcal{N}(0, \Sigma_\alpha)$ for $b = 1, ..., B = 14$ books
$\eta_q \sim \mathcal{N}(0, \sigma^2_\eta)$ for $q = 1, ..., Q = 4$ questions.
β is a 179×4 matrix of intercepts and coefficients for each strategy for each individual.

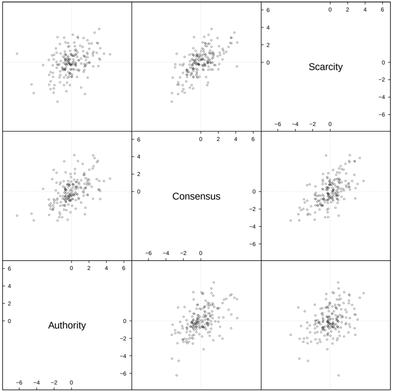
[1] Kaptein & Eckles (2012). Heterogeneity in the effects of online persuasion. *Journal of Interactive Marketing*

Results



Authority
Consensus
Scarcity
Density
Coefficient value
0.0
0.1
0.2
0.3
−5
0
5

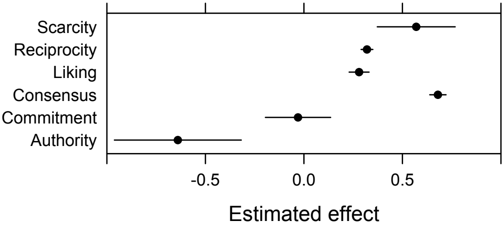
Results 2



Section 3

Persuasion Profiling

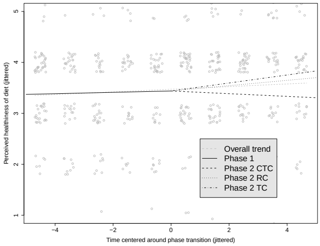
The persuasion profile[2]



[2] Kaptein, Eckles, & Davis (2011). Envisioning Persuasion Profiles. *ACM Interactions*

Using persuasion profiles for snacking[3]

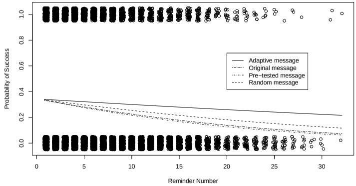
- Susceptibility measured using questionnaire
- Selection of strategies Random, Contra Tailored, or Tailored
- Expected decrease of 1 snack after 5 days
- $N = 73$



[3] Kaptein, de Ruyter, Markopoulos, & Aarts (2012). Adaptive Persuasive Systems: A Study of Tailored Persuasive Text Messages. *Transactions on Interactive Intelligent Systems*

Using persuasion profiles for email compliance[4]

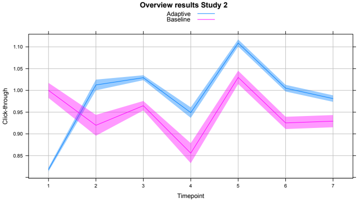
- Susceptibility estimated based on behavioural response
- Selection of strategies Adaptive, Original, Pre-tested, Random
  - More on adaptive later ...
- Large differences in success probability
- $N = 1129$



[4]Kaptein & van Halteren (2012).
Adaptive Persuasive Messaging to Increase Service Retention.
Journal of Personal and Ubiquitous Computing

Application with Booking.com[5]

- Optimising email communication for Booking.com
- 200.000+ Weekly emails
- Click through increase $> 10\%$
- Paper contains 3 more empirical validations



[5] Kaptein, Parvinen, & McFarland (2015).
Web Customization with Persuasion Profiling:
Dynamic Adaptation of Promotional Web Content on the Fly.
Under review, draft available on request.

Section 4

Recent work

Technical Challenges and Recent work

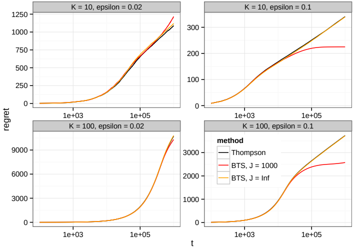
- Exploration vs Exploitation: Multi Armed Bandit Problems
  - MAB problems
  - Bootstrap Thompson sampling
  - Thompson sampling for estimation precision
- Estimation and optimization in data streams
  - SEMA (Streaming EM)
  - StreamingBandit: software
  - Lock in Feedback
- Formalizing personalization

Multi Armed Bandit

- Sequential decision making with “bandit” feedback
- Exploration vs Exploitation
- Examine policies
    - Interest in Thompson sampling

Bootstrapped Bandit[6]

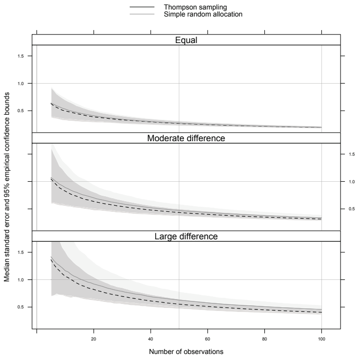
- Thompson sampling works well if posterior is known
- Not the case for complex models
- What about the (double or nothing) bootstrap distribution?
- For $1, \ldots, J$ online bootstrapped replicates



[6]Kaptein, & Eckles (2014).
Scalable Thompson Sampling with the Online Bootstrap. *arXiv*
Eckles & Kaptein (2015). Submitted.

Thompson sampling for optimal design [7]

- Thompson sampling for optimal design
- Sample for most "informative" datapoints
- Select treatments based on posterior variance estimates.



[7] Kaptein, M.C. (2014).
The Use of Thompson Sampling to Increase Estimation Precision.
Behavior Research Methods

Estimation in data streams

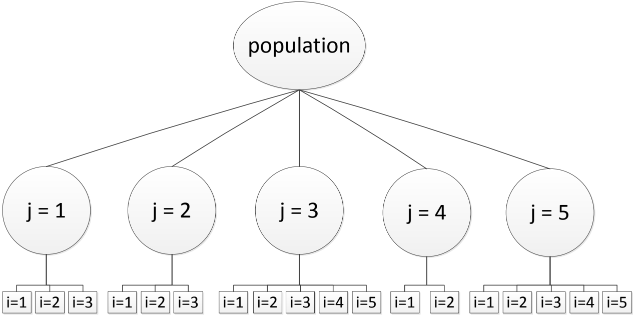


Figure: Graphical representation

Offline EM algorithm vs. SEMA[8]

Offline EM algorithm

- All data in memory
- Iterations using the entire data set
- Converges to (local) ML solution
- Refit model when new data enter

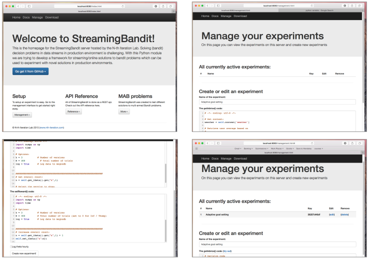
Streaming EM Approximation

- Sufficient Statistics in memory
- One iteration when a data point enters
- Converges when data stream is long enough
- Update model parameters when new data enter

[8] Ippel, L., Vermunt, J., & Kaptein, M.C. (2015) Streaming EM approximations. *Under submission.*

Streaming Bandit [9]

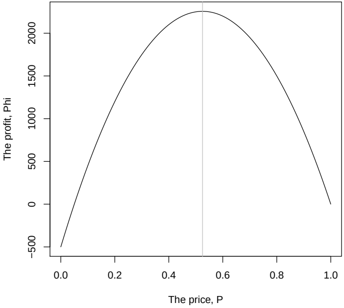
- Back end solution for streaming bandits
- Easy integration with persuasive applications
- Work in progress . . .



[9] Kaptein, M.C. & Kruijswijk, J. (2015).
Available on Github. `https://github.com/MKaptein/streamingbandit`

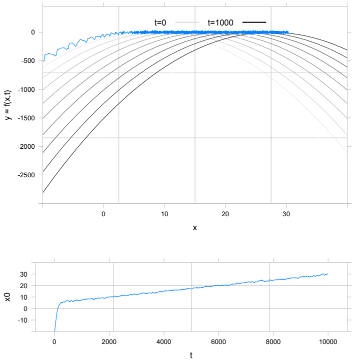
Optimization of online sales prices

- Relationship between price P and profit G.
- Exact function is however $G = f(P)$ unknown.
- We need to find the profit maximizing sales price sequentially.
- Pricing as a MAB problem
- Currently running evaluations with Santander: Pricing customer loans.



Lock in Feedback [10]

- Suppose we can vary treatments: $x(t) = x_0 + A\cos(\omega t)$ and $y = f(x)$.
- Then, we can a) multiply observed y by $\cos \omega t$, and b) integrate out the possible noise.
- By Taylor expanding y we can show this gives direct access to the derivative of (unknown) $f(x)$.
- Thus, we can use it as an update-rule for x_0.



[10] Kaptein, M.C. & Iannuzzi, D. (2015)
Lock in Feedback in Sequential Experiments. *ArXiv.*

Formalization of Personalization in Persuasive Technologies

- Features of the person and Possible treatments

$$
\begin{aligned}
x \in & \{\text{Male}, \text{Female}\} \\
a \in & \{\text{Message A}, \text{Message B}\} \\
y_i = & \mathcal{M}_g() = f(x_i, a_i)
\end{aligned}
$$

- With criterion:

$$\mathcal{C}_1 = \max \sum_{i=1}^{N} y_i = \max \sum_{i=1}^{N} f(x_i, a_i)$$

Formalization of personalized feedback

(Possible) Data generating model:

$$y_i = \mathcal{M}_G() = \beta_0 + \beta_1 a_1^* + \beta_2 x_i + \beta_3 a_i^* x_i + \epsilon$$

Personalization function:

$$a^* = \eta(a^b, x)$$

where a^b denotes the "baseline" treatment. Consider for example:

- *Non personalized:* $a_i^* = \eta_{np}(a_i, x_i) = a_i$
- *Personalized:* $a_i^* = \eta_p(a_i, x_i) = (1 - a_i)^{(1-x_i)}(a_i)^{x_i}$

Initial results [11]

Using the above formailzation *personalization* is effective if:

$$\underset{a^b}{\operatorname{argmax}} \sum_{i=1}^{N} \mathcal{M}_g(\eta_u(a_i^b), x_i)) < \underset{a^b}{\operatorname{argmax}} \sum_{i=1}^{N} \mathcal{M}_g(\eta_p(a_i^b, x_i), x_i)) \quad (2)$$

And we can state the following (trivial) results for the 2×2 model:

- Personalization is only relevant if $\beta_3 \neq 0$
- With $\eta_p(a_i, x_i) = (1 - a_i)^{(1-x_i)}(a_i)^{x_i}$ and criterion $\mathcal{C}_1$ personalization is beneficial if $\beta_3 > -\beta_1$.

[11] Kaptein, M.C. (2015) Formalizing Customization in Persuasive Technologies. *Proceedings of Persuasive 2015.*

Questions?

Maurits Kaptein
Archipelstraat 13
6524LK, Nijmegen
0621262211
maurits@mauritskaptein.com