

# Recommender System for Web Applications

Deepak Agarwal  
Yahoo! Labs

## Discussion

Leland Wilkinson  
Systat Software Inc.  
UIC CS  
NU Statistics

NISS Workshop on Computational Advertising  
Raleigh, Nov 4<sup>th</sup>, 2009

# Summary

- Explore / Exploit for multiple alternatives **outer model**
  - Each alternative is a modeled opportunity
  - k-armed bandit model governs choices
- Three classes of **inner model**
  - No personalization - pool visitors, analyze items
  - Personalization per user segment - use Reduced Rank Regression (canonical correlation model)
  - Personalization per user - use user x item Bilinear Factor Model (similar to SVD) and Hierarchical Linear Model (HLM) with covariates.

# Explore / Exploit

- Successful applications in economics, robotics, clinical trials, sequential experimental designs...
- Gittins: indices for maximizing expected discounted earnings from multi-armed bandit
- Whittle: Lagrange relaxation to allow soft bounds on constraints
- New computational efficiencies

# No Personalization Model

- Transformations?
  - Variance stabilizing (square root)
  - Anscombe (unit variance)

$$x \rightarrow 2\sqrt{x + \frac{3}{8}}$$

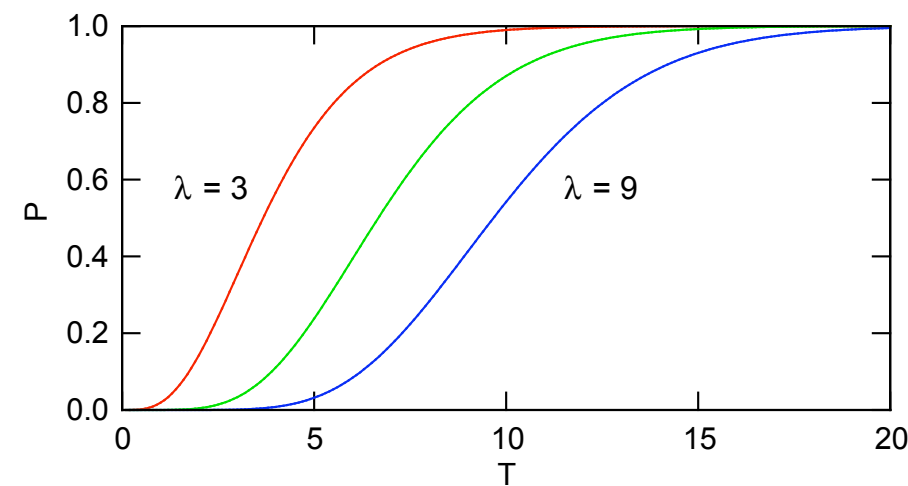
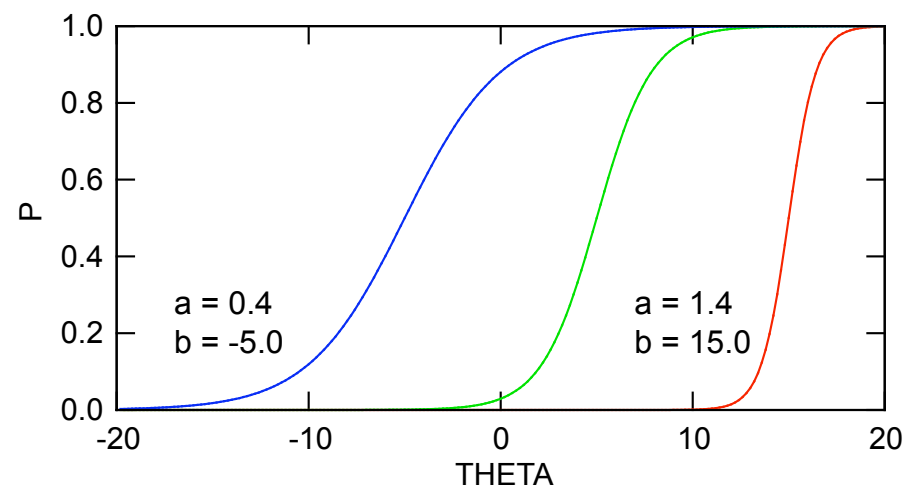
- Would this make estimation more efficient (allowing streaming data in item parameter estimation)?
- Bin widths?
  - Possibly a multi-scale problem
  - Contrast Impulsive and OCD responders
- Differencing?
  - Rate (fixed intervals) distribution Poisson, arrival time (random intervals) Exponential, Weibull, etc.
  - One way to deal with binning problem

# Personalization per User Segment Model

- Reduced Rank Regression seems appropriate for this application.
- Other approaches?
  - Least-angle, Friedman path on  $s$  parameter
  - Random projections, ...
- Can we identify exemplars?
  - Find cases near centroids of (k-means, k-medians) sequence clusters
  - Identify “outlier” response patterns / item types
  - Would assist in covariates / levels for HLM

# Personalization per User

- Cumulative logistic on left, cumulative Poisson on right.

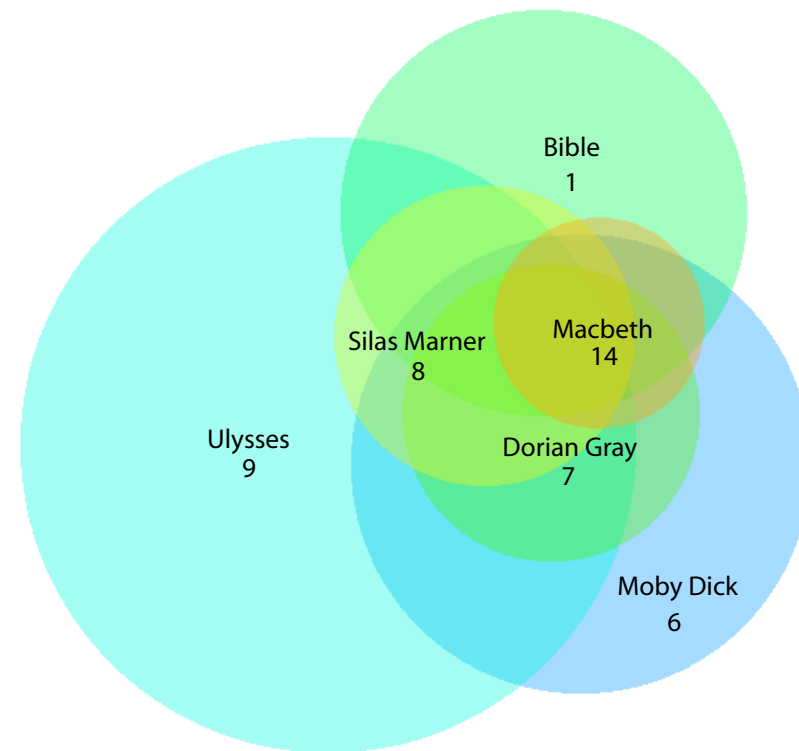


- In IRT terminology, we call X axis “difficulty” for items and “ability” for subjects. The model is of the form

$$\left\{ \begin{array}{l} f(g(X^T)) \\ g(f(X)) \end{array} \right\}$$

- In CTR terminology, we might call X axis “desirability” for items and “responsiveness” for subjects.
- Parameter estimation by Alternating Least Squares or EM. Similar to bilinear factor model.
- Tailored testing

## ■ Venn Diagrams



- Treemaps

