

# Paper Presentation – “Forecasting Repeat Sales at CDNOW: A Case Study”

Yuanlei Zhang  
April 19, 2005

## The Paper

### Forecasting Repeat Sales at CDNOW: A Case Study

PETER S. FADER  
*fader@wharton.upenn.edu*

*The Wharton School  
University of Pennsylvania  
Philadelphia, Pennsylvania 19104-6371*

BRUCE G. S. HARDIE  
*bhardie@london.edu*

*London Business School  
Regent's Park  
London NW1 4SA, United Kingdom*

INTERFACES 31: 3, Part 2 of 2, May–June 2001 (pp. S94–S107)

## To Be Covered

- Objective
- The data set
- The model
- Application
- Possible extensions

Yuanlei Zhang

2006-2-14

## Objective

- To develop a **simple** stochastic model of buyer behavior capable of forecasting medium-term **aggregate** CD purchasing by a **cohort** of new customers.

Yuanlei Zhang

2006-2-14

## Objective

- Simple
  - Implemented with common spreadsheet package like MS excel
- Aggregate
  - Using only summary data, no individual-level transaction data
- Cohort
  - Customers who first purchased at CDNOW in the first quarter (12 weeks) of 1997

Yuanlei Zhang

2006-2-14

## The Data Set

Number of CDs purchased	Week		
	1	2	...
0		1478	...
1	750	852	...
2	383	387	...
...	...	...	...
Total sales	3627	3857	...
Incremental triers	1574	1642	...
Cumulative triers	1574	3216	...

Yuanlei Zhang

2006-2-14

## The Model

- Two types of purchases in a certain week  $w$ 
  - Trial purchases ( $T_w$ )
    - The number of units purchased in week  $w$  by a customer whose trial purchase occurs in week  $w$
  - Repeat purchases ( $R_{wi}$ )
    - The number of repeat purchases made in week  $w$  by a customer whose trial purchase occurs in week  $i$  ( $i < w$ )

Yuanlei Zhang

2006-2-14

## The Model

- Modeling trial purchases
  - A shifted beta-geometric distribution is used

$$- E(T_w) = \frac{\alpha_T + \beta_T - 1}{\alpha_T - 1}$$

Yuanlei Zhang

2006-2-14

## The Model

- Modeling repeat purchases
  - A time-dependent, zero-inflated beta-geometric distribution is used

- $$E(R_{wli}) = \gamma(w-i)^\delta \frac{\beta_R}{\alpha_R - 1}$$

- The average number of units purchased decreases overtime (at a decreasing rate)

Yuanlei Zhang

2006-2-14

## The Model

- Parameter estimation
  - Model parameters:  $(\alpha_T, \beta_T, \alpha_R, \beta_R, \gamma, \delta)$
  - Estimated using MLE
  - Calculated using simple spreadsheet

Yuanlei Zhang

2006-2-14

## The Model

- Generating forecast

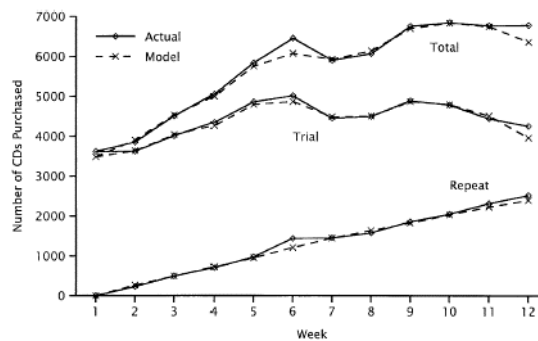
$$E(N_w) = \begin{cases} n_w E(T_w) + \sum_{i=1}^{w-1} n_i E(R_{wli}), & w \leq 12 \\ \sum_{i=1}^{12} n_i E(R_{wli}), & w > 12 \end{cases}$$

Yuanlei Zhang

2006-2-14

## Application

- Estimating total weekly sales

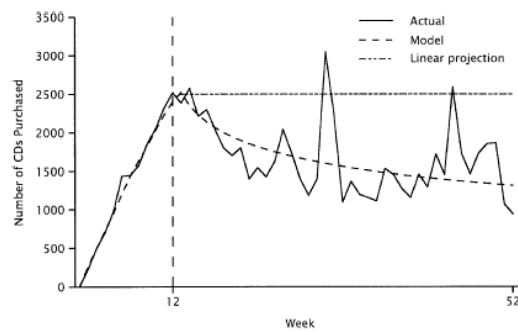


Yuanlei Zhang

2006-2-14

## Application

- Forecasting weekly repeat-sales for 40 weeks

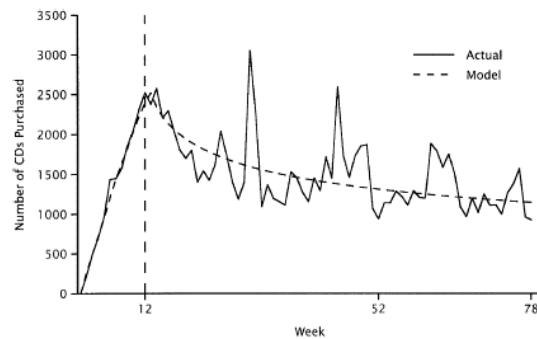


Yuanlei Zhang

2006-2-14

## Application

- Forecasting weekly repeat-sales for 66 weeks



Yuanlei Zhang

2006-2-14

## Possible Extensions

- Apply this model to other cohorts
- Develop a model for the arrival of new customers to the website
- Model more formally the dynamics of buyer behavior, using the disaggregate panel data
- Capture nonstationarity at the individual level

Yuanlei Zhang

2006-2-14

## Thank you for your patience!



Yuanlei Zhang

2006-2-14