

**“COUNTING YOUR CUSTOMERS”  
THE EASY WAY:  
AN ALTERNATIVE TO THE  
PARETO/NBD MODEL**

*P.S. Fader, B.G.S. Hardie, K.L. Lee  
Marketing Science, 2005*

*GAMES, 25/N/16*

# Структура презентации

1



2



3



4



5



- 1) Исследовательский вопрос, уникальность статьи
- 2) Данные
- 3) Модель
- 4) Результаты и выводы
- 5) Приложение к исследованию потребителей на книжном рынке

# Исследовательский вопрос

Есть ли более удобная в использовании **альтернатива** Pareto/NBD модели для расчёта **CLV** (customer lifetime value)?\*

\* сделки повторяющиеся неконтрактные



# Алгоритм поиска CLV\*

- 1) Поиск распределения функции прошлых транзакций  $f(\Theta)$ ,  $\Theta$ - индивидуальные поведенческие характеристики;
- 2) ввод предположений о распределении  $\Theta$ ;
- 3) подгонка модели под данные;
- 4) влияние  $\Theta$  на будущее ищут 2-шаговым методом:  $\Theta=f(\text{past})$ ,  $\text{future} =f(\Theta)$

\*P.S. Fader & B.G.C. Hardie, Probability Models for Customer-Base Analysis”



# Уникальность статьи: новый способ для расчет CLV

$$P(\text{«жизнь»} | x, t_x, T)$$

«Смерть»: Pareto, **BG**; «жизнь»: **NBD**

Pareto/NBD	BG/NBD
1) $X \sim \text{Pois}(\lambda t)$	1) $X \sim \text{Pois}(\lambda t)$
2) $\lambda \sim \text{Gam}(\alpha, r)$	2) $\lambda \sim \text{Gam}(\alpha, r)$
3) «Смерть» потребителей $\sim \text{Exp}(\mu)$	3) $P(\text{мгновенная «смерть» после транзакции } j) = p(p-1)^{j-1}, j \in Z$
4) $\mu \sim \text{Gam}(s, \beta)$	4) Гетерогенность $p$ задаётся $f(p   a, b) = p(p-1)^{j-1} * B(a, b)$ , где $B(a, b) = \Gamma(a) * \Gamma(b) / \Gamma(a+b)$



# Уникальность статьи: новый способ для расчет CLV

Главное **различие**: в BG/NBD  
закладывается не просто «смерть»,  
а то, что  $\forall$  потребитель может  
стать неактивным после  $\forall$   
транзакции с вероятностью  
«отсева»  $p$ .



# Данные

- покупки CD в онлайн ритейлере CDNOW, **январь 1997 – июль 1998**;
- покупатели с 1ой покупкой в 1ом квартале 1997;
- 10% выборки взято для анализа: **2357** покупателей, 78 недель, **калибровка на 39 неделях** для проверки прогноза на остальных.



# Модель VG/NBD для задачи максимизации функции L

Функция правдоподобия (L) для поведения случайного индивида

$$\begin{aligned} L(r, \alpha, a, b | X = x, t_x, T) \\ &= \frac{B(a, b+x)}{B(a, b)} \frac{\Gamma(r+x)\alpha^r}{\Gamma(r)(\alpha+T)^{r+x}} \\ &+ \delta_{x>0} \frac{B(a+1, b+x-1)}{B(a, b)} \frac{\Gamma(r+x)\alpha^r}{\Gamma(r)(\alpha+t_x)^{r+x}} \end{aligned}$$





# Модель VG/NBD для задачи максимизации функции L

$$L(r, \alpha, a, b | X = x, t_x, T) = A_1 \cdot A_2 \cdot (A_3 + \delta_{x>0} A_4)$$

$$A_1 = \frac{\Gamma(r+x)\alpha^r}{\Gamma(r)}, \quad A_2 = \frac{\Gamma(a+b)\Gamma(b+x)}{\Gamma(b)\Gamma(a+b+x)},$$

$$A_3 = \left(\frac{1}{\alpha+T}\right)^{r+x}, \quad A_4 = \left(\frac{a}{b+x-1}\right) \left(\frac{1}{\alpha+t_x}\right)^{r+x}$$

Реализация модели



# Итоговая модель

$$E(Y(t) | X=x, t_x, T, r, \alpha, a, b)$$

$$= \frac{\frac{a+b+x-1}{a-1} \left[ 1 - \left( \frac{\alpha+T}{\alpha+T+t} \right)^{r+x} {}_2F_1\left(r+x, b+x; a+b+x-1; \frac{t}{\alpha+T+t}\right) \right]}{1 + \delta_{x>0} \frac{a}{b+x-1} \left( \frac{\alpha+T}{\alpha+t_x} \right)^{r+x}}$$

$Y(t)$  – число покупок

— - гипергеометрическая функция Гаусса

— - условие покупки



# Результаты 2х моделей сопоставимы

Figure 2 Predicted Versus Actual Frequency of Repeat Transactions

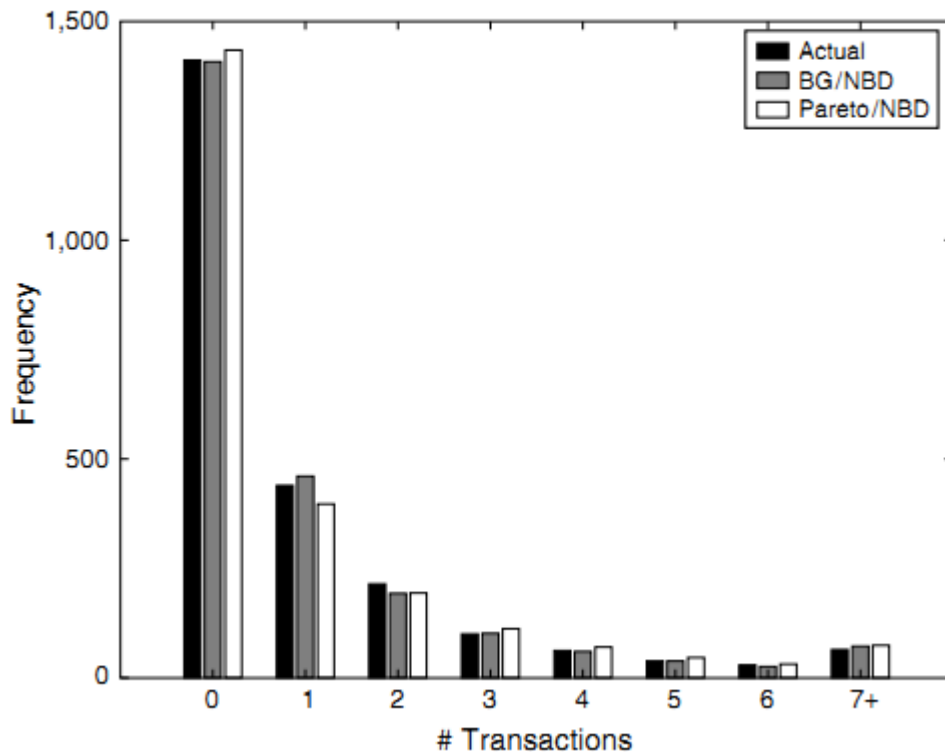


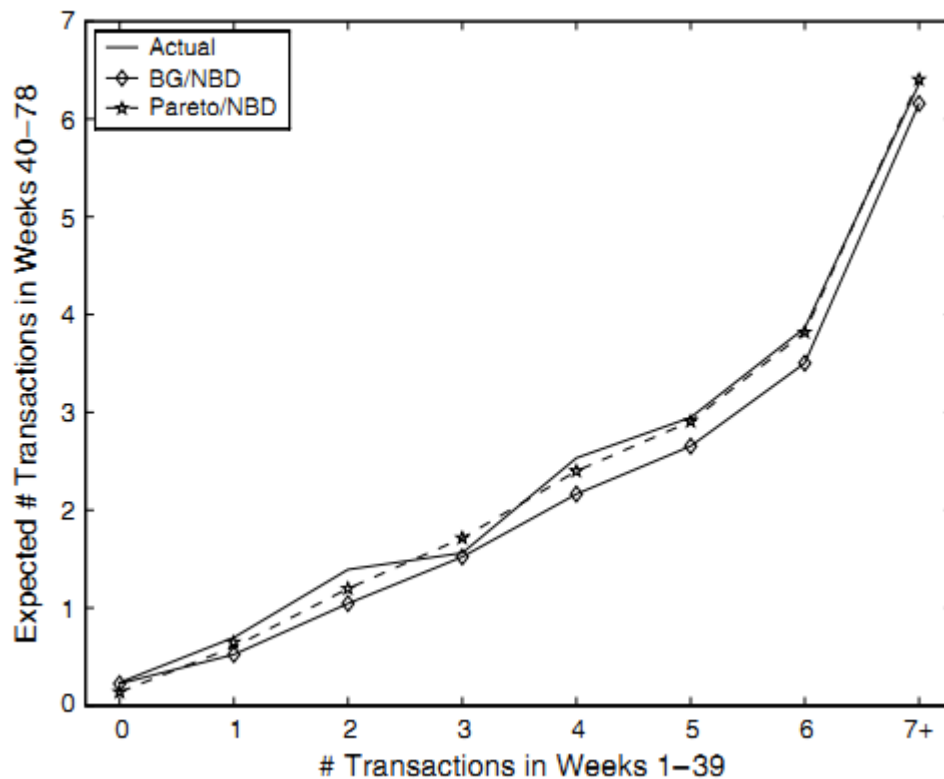
Table 2 Model Estimation Results

	BG/NBD	Pareto/NBD
$r$	0.243	0.553
$\alpha$	4.414	10.578
$a$	0.793	
$b$	2.426	
$s$		0.606
$\beta$		11.669
LL	-9,582.4	-9,595.0



# Результаты 2х моделей сопоставимы

Figure 3 Conditional Expectations



**Однако** conditional expectations для класса 0 в BG/NBD (0.23) намного ближе к реальному (0.24), чем в Pareto/NBD (0.14).

А **большая часть** данных - это **класс 0**.



# Выводы

## BG/NBD

- является **хорошей** (для прогнозирования) **альтернативой** Pareto/NBD;
- технически удобнее чем Pareto/NBD;
- лучше предсказала поведение класса 0 (критически важного).



# Приложение к исследованию потребителей на книжном рынке

## Данные

- покупки книг в ритейлере, **10.01.2013 – 05.20.2016**;
- Владельцы карточек лояльности
- Взято для анализа: когорта **07-11** месяцы **2014** года (**79** недель), **144** д/х (карточки).

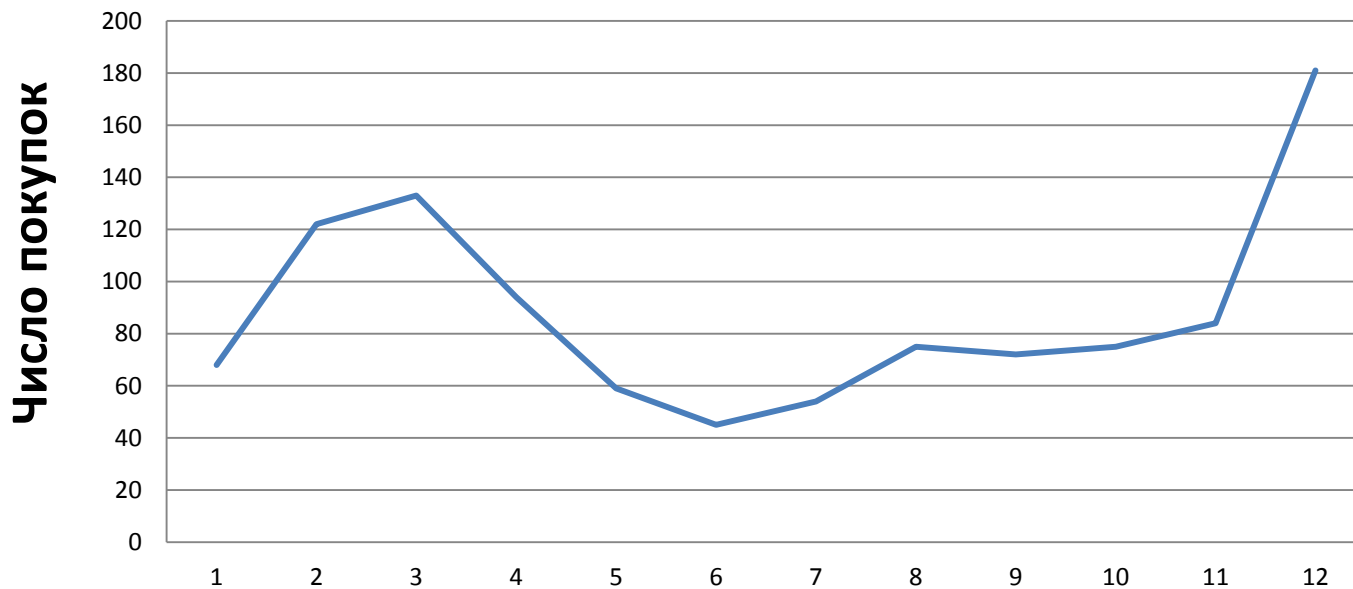
## Данные



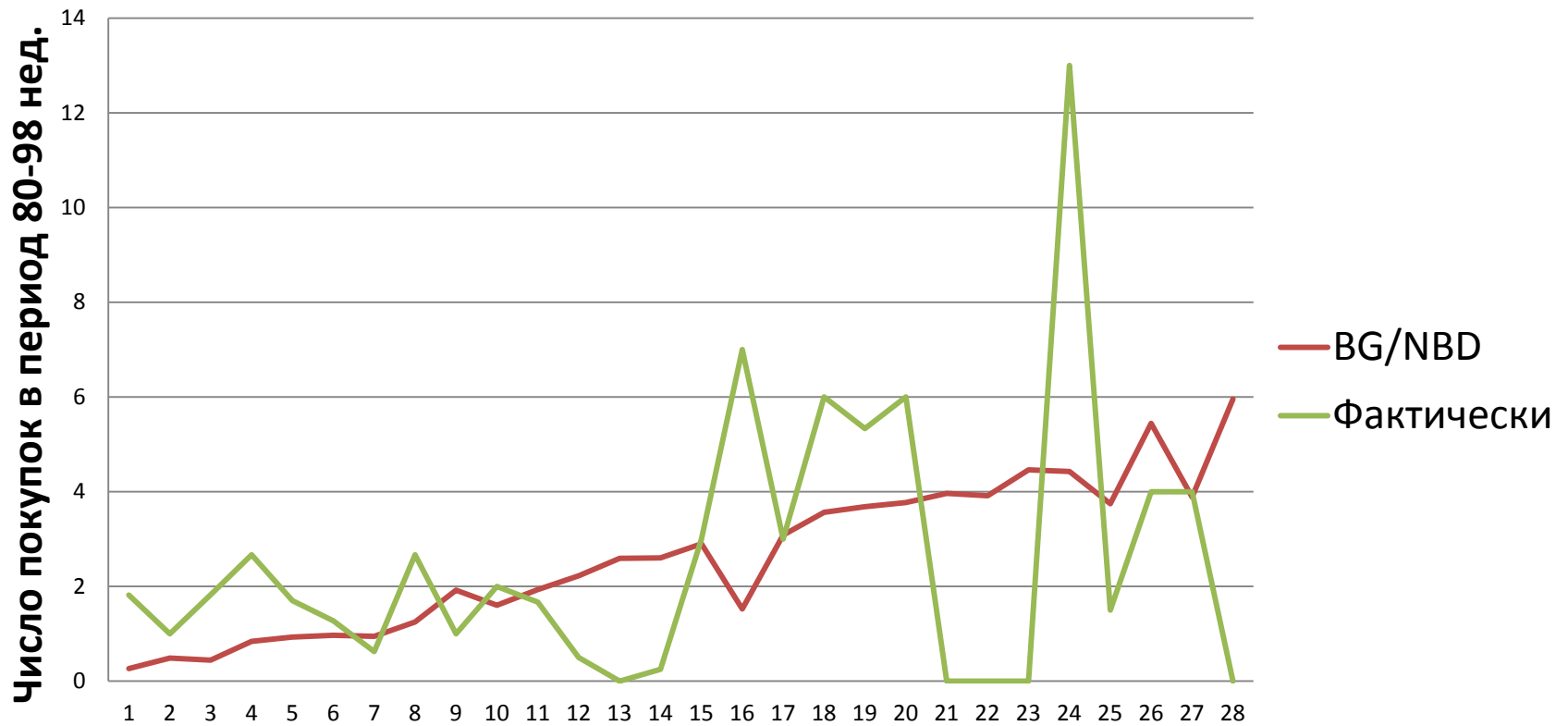
# Приложение к исследованию потребителей на книжном рынке

Обоснование когорты: 7-11 месяцы - средние значения покупок (для периода 2014-2016 гг.).

Число покупок по месяцам



# Приложение к исследованию потребителей на книжном рынке



Число покупок в период 1-79 нед.







Мулюкова Эльвира  
Mail: [mulyukovaelya@bk.ru](mailto:mulyukovaelya@bk.ru)