

MAST 6252 - Applied Probability Models

Assignment 3

1. Take the parameter estimates for the NBD, and NBD with HCNB models from the coffee creamer example that we discussed in class.

	r	α	π
NBD	0.181	0.059	
NBD with HCNB	1.226	0.215	0.551

The *penetration* of a consumer good is the proportion of the target segment that purchases at least one unit during some period of time. It is the analog to “reach” from media metrics.

- (a) What do each of the models predict will be the penetration of coffee creamer purchases after 1 period? What about after 5 periods? 26 periods?
 - (b) If you were the product manager for this particular brand of coffee creamer, and your manager asked you to estimate the maximum penetration that you could expect from this product, what *one* answer would you give? How would you justify the answer?
2. The following table, taken from a paper by Bickart and Schmittlein (1999), shows the number of surveys filled out by a sample of 1,865 Americans in 1995. Fit an NBD model to these data, discuss and interpret the parameter estimates, and forecast how many respondents are expected to complete $X = 0, 1, 2, \dots, 10$ surveys.

Number of surveys	0	1	2	3-5	6+
Number of respondents	1020	166	270	279	130

3. Posted on Blackboard is a text file, `khakichinos.txt`, that contains Internet visit data for a sample of 2,728 comScore/Media Metrix panelists who visited one particular site (with the disguised name of khakichinos.com) in July, 2014. Ignore the covariate data (the demographic information) for now.
 - (a) Fit an NBD model to the data. What do the parameter estimates tell us about the different kinds of users in the Khakichinos website user base?
 - (b) Plot the expected reach of the Khakichinos website as a function of time, from 0 to 36 months. What is the expected reach after 12 months?
4. For a typical consumer good, maintain the standard assumptions of the NBD model, where the purchase rate λ is expressed in units per month, and the mixing distribution of λ is $\text{gamma}(r, \alpha)$.
 - (a) How many units do you predict a randomly-chosen household will make in a typical month?

- (b) What is the posterior distribution of $\lambda|x, T$ for a customer who purchased x units in the last T months? (For this question, I am essentially asking you to derive the Bayes update for the NBD model. We did not do this in class, but the same principles that hold for other models will hold here as well).
- (c) The following table includes visit counts to the Khakichinos website (from Question 3) for five randomly selected users, during the past three months. Complete the table by estimating, for each user, the expected number of visits for the next month, the month after, and the two months after that. That is, if we are at the end of Month 3, what are the expected visit counts for Month 4 alone, Month 5 alone, and Months 6 and 7 combined? Use your parameter estimates from above to answer this question. To be clear, the only updating of beliefs occurs at the end of Month 3. (Hint: At each period, “the old posterior becomes the new prior”).

Months	\rightarrow	Observed counts			Expected counts		
		1	2	3	4	5	6 and 7
User	A	0	0	0			
	B	2	0	0			
	C	0	5	4			
	D	0	0	1			
	E	6	5	4			