Coding Assignment 2: Generative Indian Buffet Process

Susan Cherry

In this assignment, I'm asked to code the generative Indian Buffet Process. First, a quick review of the Indian Buffet Process: N customer enter the restaurant one after another. Each customer encounters a buffet of infinitely many dishes. The first customer moves down the line and stops after sampling Poisson(α) dishes. The ith customer moves along the buffet and samples each dish according to how many people have tried it already, $\frac{m_k}{i}$, where m_k is the number of previous customers who have sampled the dish. At the end of all of the previously sampled dishes, the ith customer will trie Poisson($\frac{\alpha}{i}$) more dishes.

See my attached code for the implementation of the IBP. I represent the IBP with binary matrices. The rows represent customers and the columns are dishes. If customer i has sampled dish j, the ith, jth point of the matrix is black. Otherwise it is white.

I vary the α parameter to investigate how my results change. α controls the number of new dishes that a customer samples, so I expect the number of dishes to increase as α increases. Note that while I vary α I keep the number of customers fixed at N=30. Below is the binary matrix for $\alpha = 1$. As the figure shows, there are only 6 dishes, meaning that most customers do not choose to sample any new dishes.

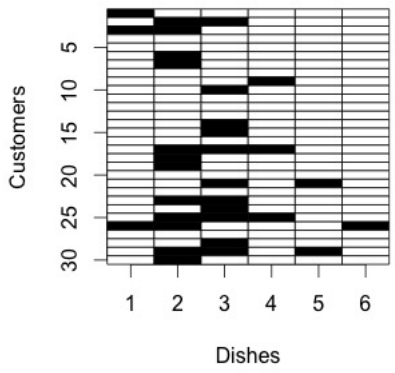

Figure 1: IBP with $\alpha = 1$ and N=30

Next, I increase α to $\alpha = 10$. Figure 2 shows the resulting binary matrix. Now there are many more dishes-nearly 40. As expected, nearly all customers sample the first dishes, but as we go down the line, fewer customers try the newer dishes.

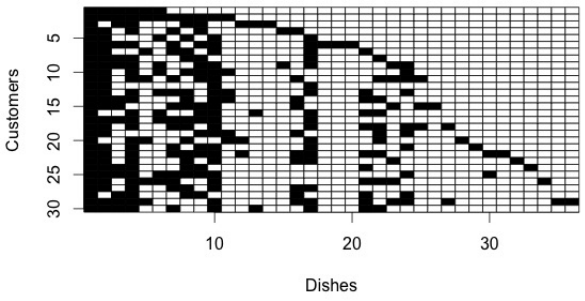

Figure 2: IBP with $\alpha = 10$ and N=30

Finally, I increase α to $\alpha = 20$. Figure 3 shows the results. The number of dishes has increased to nearly 70, as we would expect from increasing α. Again, most customers try the dishes at the beginning of the line and fewer try the ones farther down.

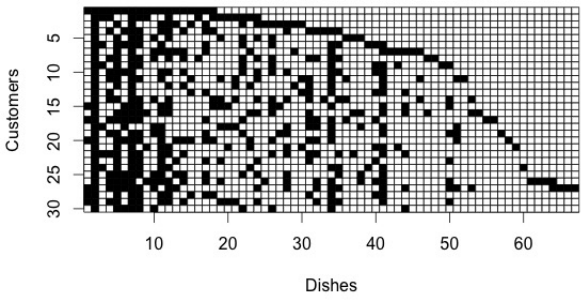

Figure 3: IBP with $\alpha = 20$ and N=30

Next I investigate how my results change as I change the number of customers. I would expect the number of dishes to increase as the number of customers increases. For these results, I fix α at $\alpha = 10$. Figure 4 shows the results for $N = 5$. In this case, a total of 20 dishes are sampled.

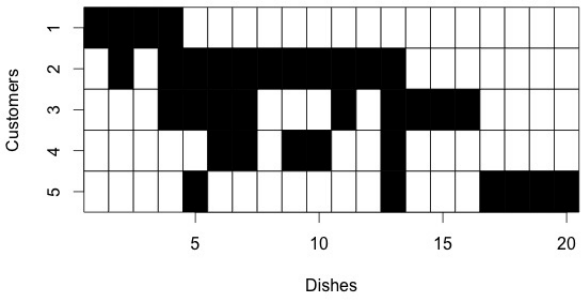

Figure 4: IBP with $\alpha = 10$ and N=5

Next, I increase the number of customers to 50. Figure 5 shows that the number of dishes increases dramatically to nearly 60.

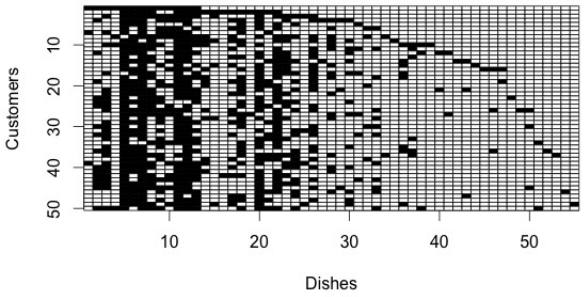

Figure 5: IBP with $\alpha = 50$ and N=5

Finally, I increase the number of customers even further to 100. As Figure 6 shows, the number of dishes increases again to 70. In summary, I find that the number of dishes sampled in an Indian Buffet Processes is increasing in both α and in the number of customers that arrive at the restaurant.

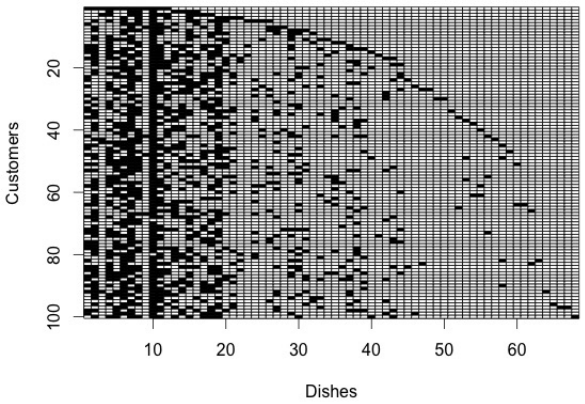

Figure 6: IBP with $\alpha = 100$ and N=5