Hidden Markov Model of Portfolio Credit Risk*

Saran Ahuja, Sarita Bunsupha, Kate Tan, Xiaokan Tong, Vicky Wen
Stanford University

June 7, 2011

Abstract

In this paper, we model the default arrival using a Hidden Markov Model. This model is motivated by actual default data from 1970 - 2006 which suggests periods of different intensities. We apply the Expectation Maximization Algorithm to estimate the parameters involved, and the forward-backward Viterbi algorithm to decode the most likely path of hidden intensities. In-sample and out-of-sample tests are performed to evaluate our fitted model.

1 Introduction

The market for Credit Default Swaps (CDS) and Collaterized Debt Obligation (CDO) has been growing significantly over the past few decades. As a result, there is a significant need for a method to price these instruments which requires modeling the default rates. There are many ways to model defaults but, recently, the common practice is to use intensity based models. Using an intensity based model requires specifying the dynamics of the intensity. A simple approach is modeling the intensity to be constant. More sophisticated approaches include modeling the intensity as a stochastic process, such as a Hawkes process or Feller diffusion process. In this report, we propose another model for the intensity–a hidden Markov chain. This model is simpler than the stochastic models but still captures the high/low regime feature exhibited by the data.

This report is organized as follows: Section 2 describes the data and the hidden Markov model that we will use throughout the paper. Section 3 describes the methods used to estimate the parameters and find most likely hidden states. Section 4 gives estimation results for various models. Section 5 describes the performance analysis of our model. Finally, Section 6 concludes our findings.

*This paper was done as a final project for the course MS&E 347: Credit Risk: Modeling and Management at Stanford University taught by Prof. Kay Giesecke

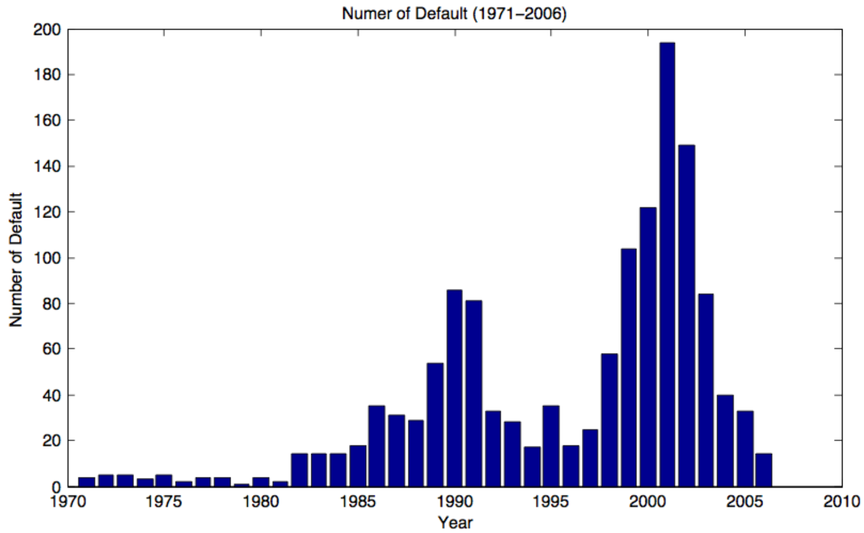

Figure 1: Plot of number of defaults from 1970 to 2006

2 Hidden Markov Model

2.1 Motivation

Consider the following Data on economy-wide, industrial and financial default timing from Moodys Default Risk Service. The sample period is January 1970 to October 2006. The data observed are the dates when each default occurs. There are 1,389 defaults on 919 unique days–multiple defaults occurred on certain days and the time of day is not specified in the data. The data consists of senior rated corporates (approximately 6,048 firms of which 3,215 are investment grade). A default is a credit event in any of Moodys default categories. Roughly 51% of the defaults are due to missed interest payments and 25% are due to Chapter 11. Approximately 4 of the defaults were investment grade. The plot of number of defaults is given in figure 1.

From figure 1, one can easily see that the rate of default is not constant across time. There is a period where the intensity is low (1970 - 1980) and a period where the intensity is high (1998 - 2004). Motivated by this observation, we will model the underlying intensity as an unobservable Markov chain. The detail of our model is given in the next section.

2.2 Model

In this section, we describe the mathematical framework for the Hidden Markov model that we will be using. We assumed that the observed default arrivals are realizations of a counting process whose intensity is a function of a Markov chain that cannot be observed. Observing the default arrival, we can use Bayes' rule to find the probability of the intensity being in each state conditional on the observation.

For simplicity, we will work with a discrete model. One reason for this decision is that we would l need to discretize the time step and work with a discrete time model anyways when estimating the parameters and forecasting. Thus, it is reasonable to start with the discrete model. Another reason is that using a discrete model, we can alter the default distribution at each step to other distributions such as Binomal or Negative Binomial.

Suppose we observed number of defaults from time 0 to T. Let N be the number of discretization steps, $N\Delta t = T$, and $t_k = k\Delta t$. We will let Δt be one week or one day. Let X_n denote the intensity (or probability of failure in the case of the Binomial distribution) at time t_n. We assume that X_n is a Markov Chain with finite state space $\{\lambda_1, \ldots, \lambda_{nS}\}$ or $\{q_1, \ldots, q_{nS}\}$, where nS is either two, three, or four states, with transition matrix P. Let Z_n denote the number of defaults at time t_n and f_j be the probability mass function of Z_n given that X_n is in state j. That is,

$$\mathbb{P}[Z_n = k | X_n = \lambda_j] = f_j(k)$$

For example, we can let Z_n be distributed as a Poisson random variable where X_n determines the intensity, i.e.

$$f_j(k) = \frac{e^{-\lambda_j}\lambda_j^k}{k!}$$

Alternatively, we can let Z_n be distributed as a Binomial random variable, i.e.

$$f_j(k) = \binom{N_F}{k}\left(\frac{q_j}{N_F}\right)^k\left(1 - \frac{q_j}{N_F}\right)^{N_F - k}$$

where N_F denotes the number of surviving firms, which we will take to be constant $N_F = 6048$. The main feature of our model is that X_n is unobservable. We will need to compute the conditional distribution of X_n given the observations $Z_1, \ldots, Z_n$. In addition, we also do not know neither the state X_n can take nor the probability transition matrix. As we shall see in the next section, all these parameters will have to be estimated using what is called the EM algorithm.

3 Parameter Estimation Method

In this section, we describe two main tools widely used to solve problems relating to Hidden Markov Models (HMM). The first tool is the Expectation Maximization (EM) algorithm and is used to estimate parameters of the underlying Markov chain. The second tool is called the Viterbi algorithm and is used to find the most likely sample path of unobservable Markov chain. The details of how to implement these two methods follows below,

3.1 EM Algorithm

The EM algorithm is based on the maximum likelihood idea. It is an iterative method that alternates between performing an expectation (E) step, which computes the expectation of the log-likelihood using the current estimates, and a maximization (M) step, which computes parameters by maximizing the expected log-likelihood computed in the (E) step. Please refer to [1] for details on how to implement EM algorithm.

3.2 Viterbi Algorithm

The Viterbi algorithm is a forward-backward programming algorithm for finding the most likely sequence of hidden states. First, we compute the maximal probability of being in state i at time n. This step is done forward in time. Then once we have computed these probabilities for every state and time, we find the optimal path by finding the argument that maximizes the values we found in the first step. This step is done backward in time. For details of how to implement the Viterbi Algorithm, please refer to [2].

4 Model Calibration

In this section, we present the results of applying the EM and Viterbi algorithms described in Section (3.1) and Section (3.2) to the data described in Section (2.1). We varied the model from two to four states as well as the distribution from Poisson and Binomial distribution. Note that the EM algorithm only gives us the local maximum of the log-likelihood function, hence it depends upon our choice of initial guess. By analyzing the data, we determined several initial guesses. Then we applied both algorithms and selected the one that gave the highest log-likelihood.

4.1 Poisson Distribution

In this section, we assume

$$f_j(k) = \frac{e^{-\lambda_j} \lambda_j^k}{k!}$$

The result of applying the EM algorithm is as follows,

Two States	Initial Parameters	Final Estimated Parameters
Initial Probability	[0.5 0.5]	[1 0]
Intensity	[0.15 2.25]	[0.2209 2.1564]
Transition Matrix	$\begin{bmatrix} 0.33 & 0.33 \\ 0.33 & 0.33 \end{bmatrix}$	$\begin{bmatrix} 0.9787 & 0.0166 \\ 0.0710 & 0.9400 \end{bmatrix}$

Table 1: Estimated parameters using the EM algorithm for Poisson/two state model

Three States	Initial Parameters	Final Estimated Parameters
Initial Probability	[0.5 0.5 0.5]	[1 0 0]
Intensity	[0.01 0.15 0.25]	[0.0813 0.6452 3.1661]
Transition Matrix	$\begin{bmatrix} 0.33 & 0.33 & 0.33 \\ 0.33 & 0.33 & 0.33 \\ 0.33 & 0.33 & 0.33 \end{bmatrix}$	$\begin{bmatrix} 0.9855 & 0.0092 & 0.0000 \\ 0.0092 & 0.9568 & 0.0299 \\ 0.0127 & 0.1335 & 0.8816 \end{bmatrix}$

Table 2: Estimated parameters using the EM algorithm for Poisson/three state Model

Four States	Initial Parameters	Final Estimated Parameters
Initial Probability	[0.5 0.5 0.5 0.5]	[1 0 0 0]
Intensity	[0.07 0.7 2.3 3.75]	[0.0639 0.4924 1.1325 4.1806]
Transition Matrix	$\begin{bmatrix} 0.33 & 0.33 & 0.33 & 0.33 \\ 0.33 & 0.33 & 0.33 & 0.33 \\ 0.33 & 0.33 & 0.33 & 0.33 \\ 0.33 & 0.33 & 0.33 & 0.33 \end{bmatrix}$	$\begin{bmatrix} 0.9943 & 0.0020 & 0.0000 & 0.0031 \\ 0.0000 & 0.9950 & 0.0000 & 0.0037 \\ 0.0000 & 0.0000 & 0.3142 & 0.6888 \\ 0.0163 & 0.0409 & 0.9110 & 0.0313 \end{bmatrix}$

Table 3: Estimated parameters using the EM algorithm for Poisson/four state model

Please refer to the Appendix for figures showing the convergence of the EM algorithm and most likely hidden intensities.

4.2 Binomial Distribution

In this section, we assume

$$f_j(k) = \binom{N_F}{k} \left(\frac{q_j}{N_F}\right)^k \left(1 - \frac{q_j}{N_F}\right)^{N_F - k}$$

Two States	Initial Parameters	Final Estimated Parameters
Initial Probability	[0.5 0.5]	[1 0]
Intensity	[0.15 2.25]	[0.2257 2.1708]
Transition Matrix	$\begin{bmatrix} 0.33 & 0.33 \\ 0.33 & 0.33 \end{bmatrix}$	$\begin{bmatrix} 0.9794 & 0.0159 \\ 0.0687 & 0.9426 \end{bmatrix}$

Table 4: Estimated parameters using the EM algorithm for Binomial/two states model

Three States	Initial Parameters	Final Estimated Parameters
Initial Probability	[0.5 0.5 0.5]	[1 0 0]
Intensity	[0.01 0.15 0.25]	[0.0814 0.6457 3.1729
Transition Matrix	$\begin{bmatrix} 0.33 & 0.33 & 0.33 \\ 0.33 & 0.33 & 0.33 \\ 0.33 & 0.33 & 0.33 \end{bmatrix}$	$\begin{bmatrix} 0.9854 & 0.0093 & 0.0000 \\ 0.0093 & 0.9563 & 0.0304 \\ 0.0127 & 0.1356 & 0.8794 \end{bmatrix}$

Table 5: Estimated parameters using the EM algorithm for Binomial/three state Model

Four States	Initial Parameters	Final Estimated Parameters
Initial Probability	[0.5 0.5 0.5 0.5]	[1 0 0 0]
Intensity	[0.07 0.7 2.3 3.75]	[0.0639 0.4924 1.1326 4.1815]
Transition Matrix	$\begin{bmatrix} 0.33 & 0.33 & 0.33 & 0.33 \\ 0.33 & 0.33 & 0.33 & 0.33 \\ 0.33 & 0.33 & 0.33 & 0.33 \\ 0.33 & 0.33 & 0.33 & 0.33 \end{bmatrix}$	$\begin{bmatrix} 0.9943 & 0.0020 & 0.0000 & 0.0031 \\ 0.0000 & 0.9950 & 0.0000 & 0.0037 \\ 0.0000 & 0.0000 & 0.3146 & 0.6884 \\ 0.0163 & 0.0409 & 0.9110 & 0.0312 \end{bmatrix}$

Table 6: Estimated parameters using the EM algorithm for Binomial/four states model

Please refer to the Appendix for figures showing the convergence of the EM algorithm and most likely hidden intensities.

5 Performance Analysis

From the results in the prior section, one can see that the Binomial and Poisson give very similar results. As a result, we will only consider the Poisson model from this point onward. In this section, we will analyze the fitted model by doing in-sample tests in Section 5.1 . Then, after applying in-sample tests for two, three and four state models, we select the best model to perform forecasting tests on Section 5.2.

5.1 In-sample Analysis

In this section, we perform an in-sample test to see if our estimated model fits well with the data. The idea is to apply a time-change method to get the time-rescaling inter arrival time. That is, we need to keep track of

$$\hat{\lambda}(n) := \mathbb{E}[X_n | Z_1, \ldots, Z_n] = \sum_j \pi(n, j)\lambda_j$$

where

$$\pi(k, j) = \mathbb{P}[X_k = \lambda_j | Z_1, \ldots, Z_k]$$

These rescaled arrival times should then follow standard exponential random variables. We use a QQ plot to measure the goodness of fit of the data to the standard exponential distribution, Kolmogorov-Smirrnov (KS) test and Chi-square test to test if it follows iid standard exponential distribution, and the Box-Ljung test to test the independence of inter-arrival events. Note that these tests are only valid in the case of the Poisson distribution. The results for the two, three, and four state models are as follows,

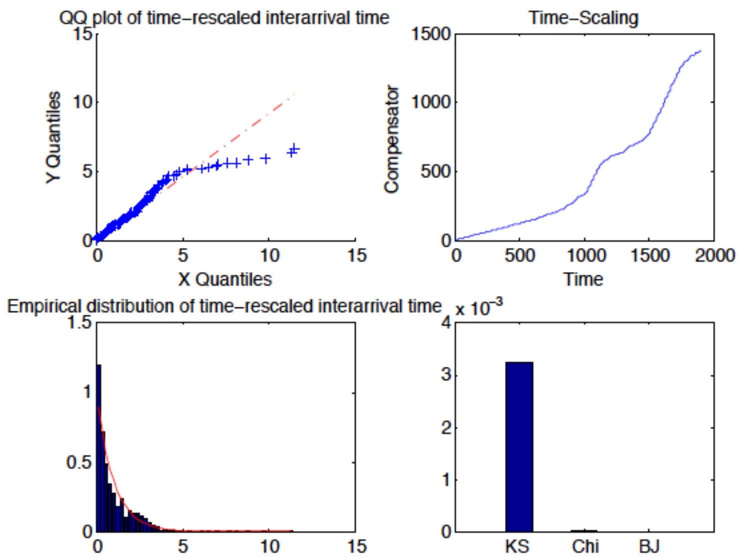

Figure 2: Plot of in-sample test under Poisson/two state model

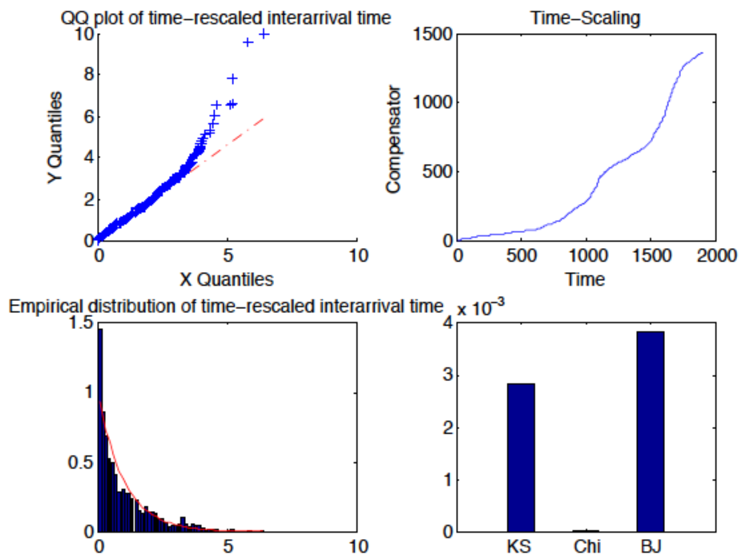

Figure 3: Plot of in-sample test under Poisson/three state model

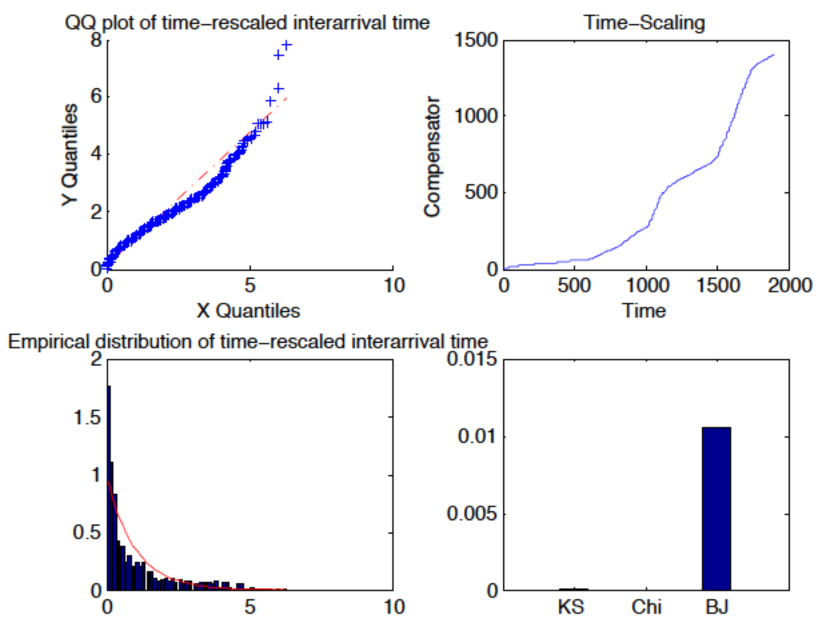

Figure 4: Plot of in-sample test under Poisson/four state model

From Figure 2 to 4, one can easily see that the four state model yields the best fit. In the case of the two and three state models, the rescaled inter arrival times deviate from the standard exponential distribution very quickly. In contrast, the four state model fits the standard exponential almost perfectly. This observation is in line with the significant improvement in all three tests (KS, Chi-square, BL). As a result, we choose this model for our out-of-sample test.

5.2 Out-of-sample Analysis

To predict the number of defaults in year Y, we first apply the EM algorithm based on the historical data from 1970 to year $Y-1$ to estimate the parameters. To perform the forecasting, we follow the method explained in class. That is, we generate a sequence of standard exponential distribution and obtain the rescaled inter arrival times. Then we compute the actual inter arrival times and count the number of defaults, Z_k. We take this new information and find the expected intensity in the next step. This can be done as follows; We let

$$q(k,j) = \mathbb{P}[X_k = \lambda_j, Z_i = z_i, 1 \leq i \leq k-1]$$

and

$$\pi(k,j) = \mathbb{P}[X_k = \lambda_j | Z_i = z_i, 1 \leq i \leq k-1] = \frac{q(k,j)}{\sum_i q(k,i)}$$

Thus, we only need to update $p(n,i)$. Note that

$$\begin{aligned} q(n+1,i) &= \mathbb{P}[X_{n+1} = \lambda_i, Z_l = z_l, 1 \leq l \leq n] \\ &= \sum_j \mathbb{P}[X_{n+1} = \lambda_i, X_n = \lambda_j, Z_l = z_l, 1 \leq l \leq n] \\ &= \sum_j \mathbb{P}[X_{n+1} = \lambda_i | X_n = \lambda_j, Z_l = z_l, 1 \leq l \leq n]\mathbb{P}[X_n = \lambda_j, Z_l = z_l, 1 \leq l \leq n] \\ &= \sum_j p_{ji}\mathbb{P}[Z_n = z_n | X_n = \lambda_j, Z_l = z_l, 1 \leq l \leq n-1]\mathbb{P}[X_n = \lambda_j, Z_l = z_l, 1 \leq l \leq n-1] \\ &= \sum_j p_{ji} f_j(z_n) q(n,j) \end{aligned}$$

The results of the out-of-sample test are

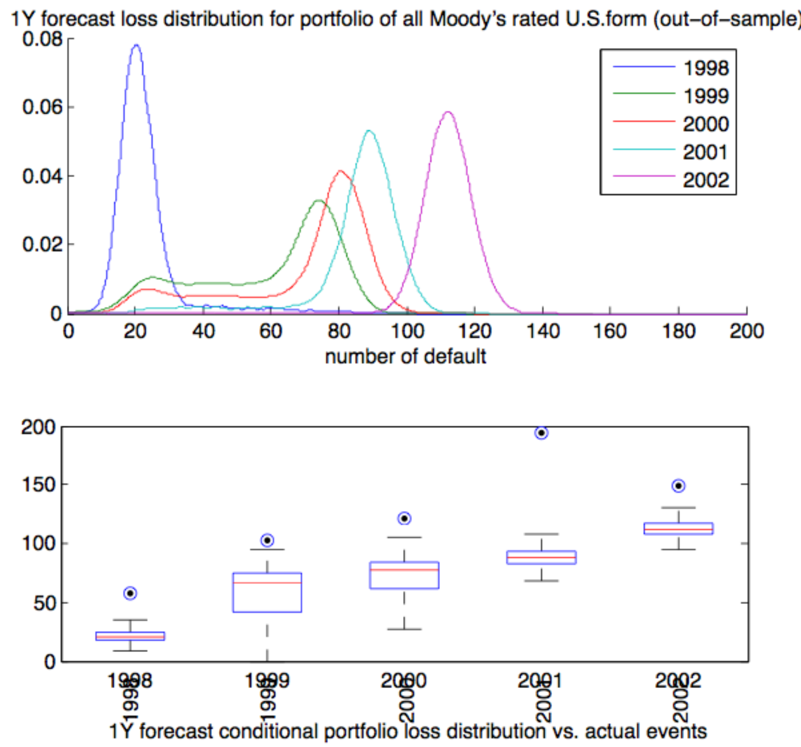

Figure 5: Plot of out-of-sample test

6 Conclusion

In this project, we model the default arrival using a discrete time Hidden Markov Model (HMM). By applying the EM algorithm, we estimated the transition probability and intensity in each state. Also, we applied the forward-backward Viterbi algorithm to find the most likely hidden states.

After fitting our model to the data, the results are very similar for the Poisson and Binomial model. As a result, we decided to work with the Poisson and perform in-sample tests. The result is as expected; the in-sample test improved as we increased the number of states. The QQ plot of rescaled inter-arrival times from Poisson/four state model fits almost perfectly with the standard exponential distribution.

Finally, we selected the best model and performed a forecasting test (out-of-sample test). However, our model does not perform well in regards to forecasting. It failed to catch the growing number of defaults during the years 1998 - 2002. This is not surprising considering the structure of the data and how our model is constructed. We assumed a fixed number of states and using data from year 1970 to year $Y-1$ to forecast the number of defaults in year Y. The data used to estimate the parameters has not seen the enhanced risk state before and, as a result, all parameters involved are underestimated. To fix this problem, it is important that the data used for estimation include several periods of economic crisis where the number of defaults is extremely high.

Appendix A : Figure

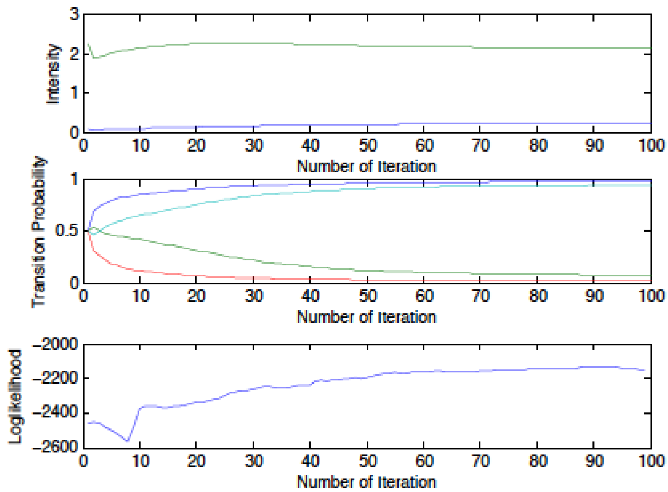

Figure 6: Plot of convergence of parameter estimators from EM algorithm under Poisson/two states model

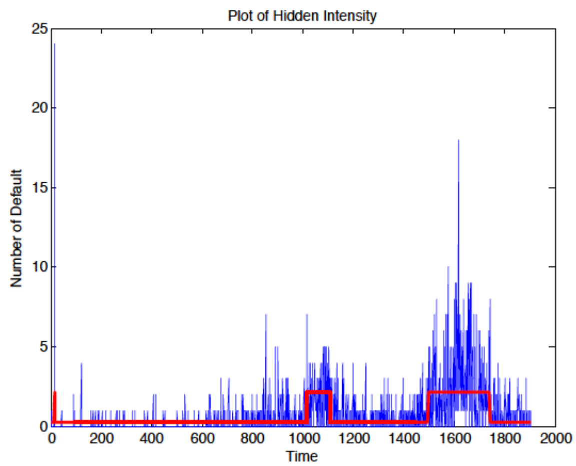

Figure 7: Plot of optimal hidden states under Poisson/two states model

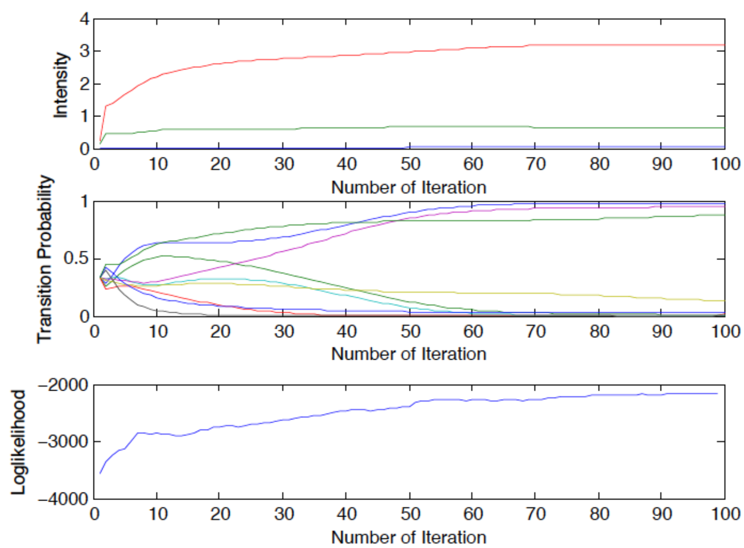

Figure 8: Plot of convergence of parameter estimators from EM algorithm under Poisson/three states model

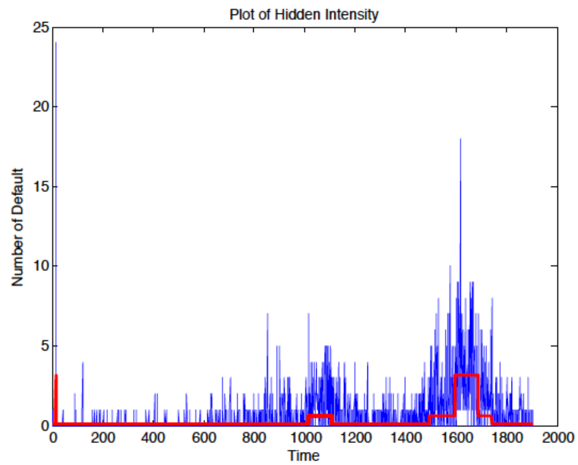

Figure 9: Plot of optimal hidden states under Poisson/three states model

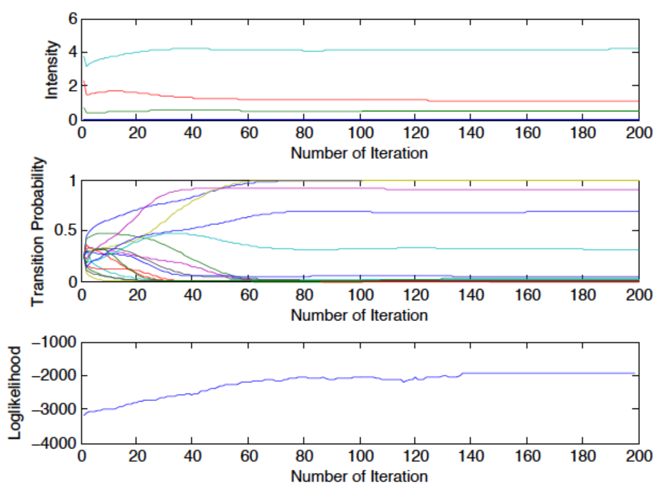

Figure 10: Plot of convergence of parameter estimators from EM algorithm under Poisson/four states model

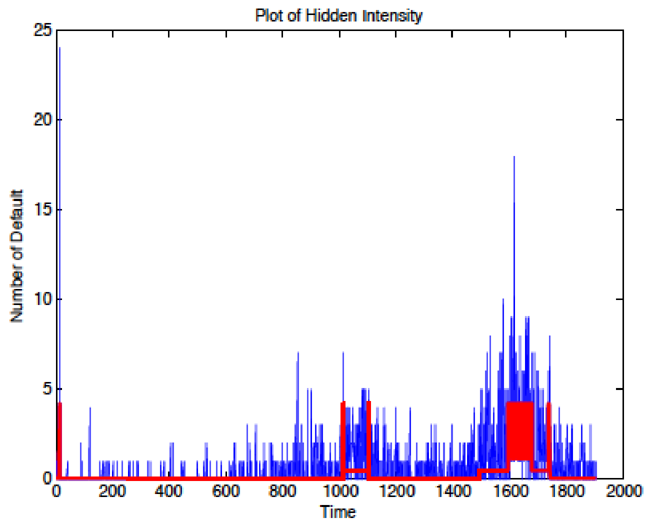

Figure 11: Plot of optimal hidden states under Poisson/four states model

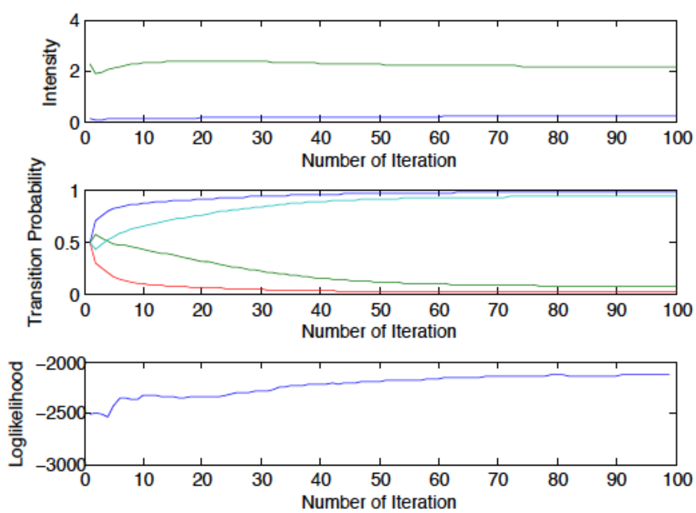

Figure 12: Plot of convergence of parameter estimators from EM algorithm under Binomial/two states model

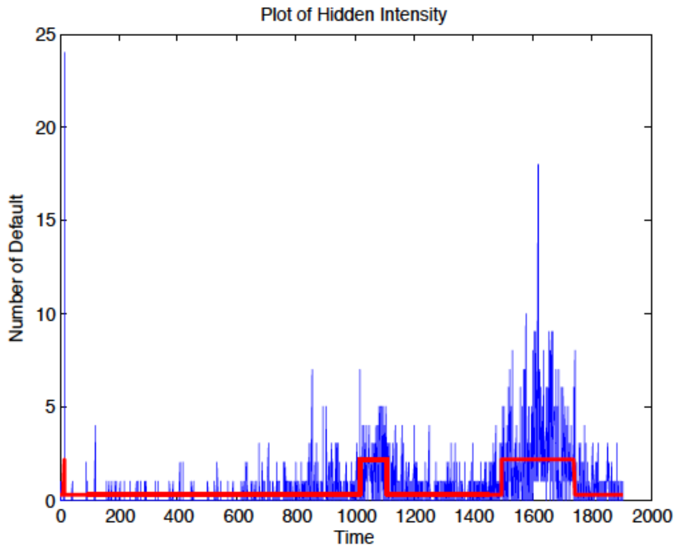

Figure 13: Plot of optimal hidden states under Binomial/two states model

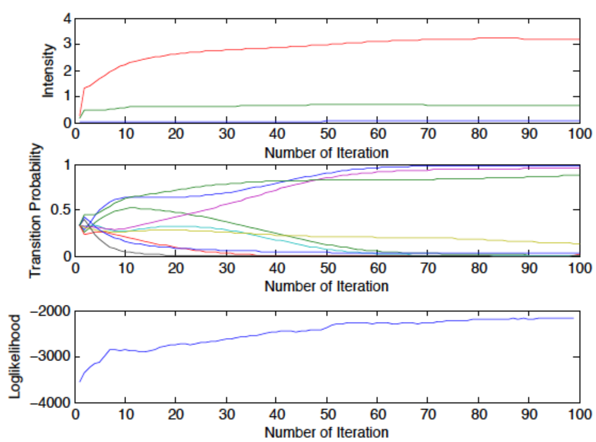

Figure 14: Plot of convergence of parameter estimators from EM algorithm under Binomial/three states model

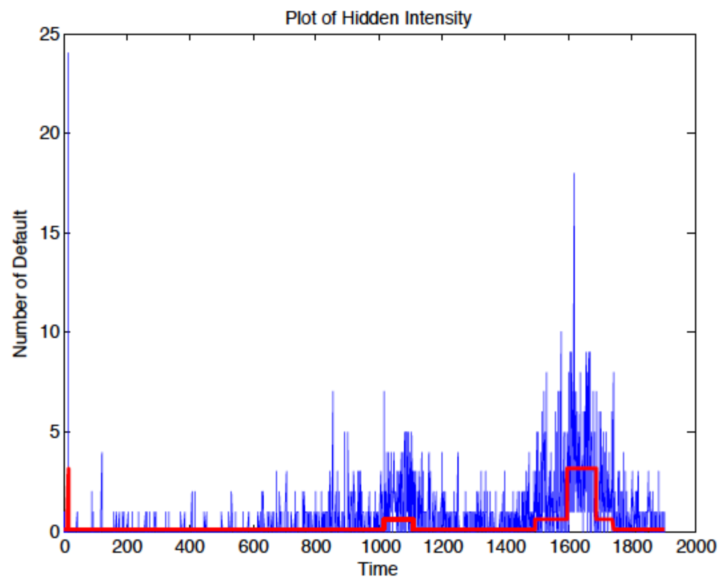

Figure 15: Plot of optimal hidden states under Binomial/three states model

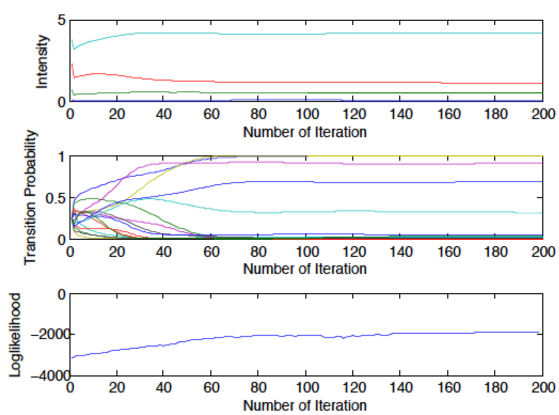

Figure 16: Plot of convergence of parameter estimators from EM algorithm under Binomial/four states model

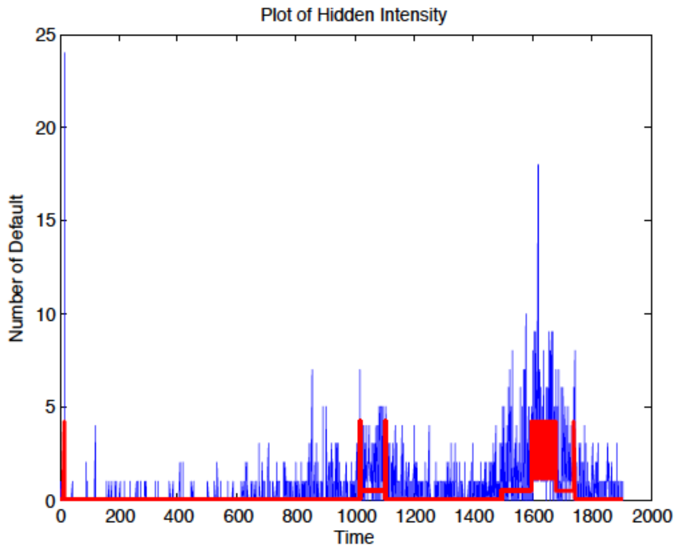

Figure 17: Plot of optimal hidden states under Poisson/four states model

Appendix B : Group Members Distribution

- Saran Ahuja: model set up, Poisson model, in-sample test, out-of-sample test, report, presentation.
- Sarita Bunsupha: model set up, Poisson model, out-of-sample test, presentation.
- Kate Tan: Binomial model, out-of-sample test, report, presentation.
- Nathan Tong: in-sample test.
- Vicky Wen: Binomial model.

References

[1] David Vere-Jones Daryl J. Daley, *Introduction to the theory of point processes: General theory and structure*, vol. 2, Springer, 2007.

[2] B.H. Juang L.R. Rabiner, *An introduction to hidden markov models*, IEEE ASSP Magaine (1986).