Stochastic Modeling and Applications
Vol. 24 No. 1 (January-June, 2020)

Received: 13th January 2020 Revised: 07th February 2020 Accepted: 12th February 2019

MODELLING TIME SERIES OF COUNT DATA USING HIDDEN MARKOV MODELS

AMBILY JOSE[1,*], TUNNY SEBASTIAN[2] AND SEBASTIAN GEORGE[1]

Abstract. Poisson distribution is normally used to model count data. However, in practical situations, the equidispersion feature of Poisson tends to go wrong. In such cases, mixture models are usually used to fit the data. But serial dependence in the observations does not agree with the mixture models and hence, they assume independence. The Hidden Markov Model (HMM) comes in handy in such a scenario as it handles both overdispersion (variance greater than mean) and serial dependence in the data. HMM is an extension of Markov models and it can be considered as a generalization of mixture model. In this study, we generate two sets of count data - (i) independent and overdispersed and (ii) serially-dependent and overdispersed and conduct a simulation study for proving the performance of HMM. For the study, the Poisson-Hidden Markov (P-HMM) Model is applied to a real data on the monthly cases of Hepatitis-B in the state of Kerala in South India, which is overdispersed and serially dependent. Viterbi algorithm is used to obtain the best estimate of the state sequence. By applying Akaike information criterion (AIC) and Bayesian information criterion (BIC) model selection criteria, we could find that the prevalence of Hepatitis-B can be modelled by 5-state P-HMM.

1. Introduction

Time series of disease count data is serially correlated and numerous models have been suggested for their analysis. Time series count, in general, are molded by autoregressive and moving average (ARMA) model [1], which is based on the normal distribution. Disease count in unit area is usually modelled by Poisson distribution. The equidispersion feature is one of the important characteristics of Poisson distribution. However, in many practical cases, either mean will be greater than variance or vice-versa, making Poisson assumption wrong. Real-life count data, including epidemiological surveillance stats, are frequently characterized by overdispersion.

Mixture models are used to deal with overdispersed observations which are usually a bimodal or a multimodal distribution. But serial dependence in the observations does not agree for the mixture models and so it assumes independence. When it comes to practical applications, the independence assumption usually fails. A method for permitting serial dependence in the observations is to assume that parameter process is consecutively dependent. A suitable technique to do so is to assume that it is a Markov chain (MC). The resulting model for the observations is called an HMM, which can accommodate both overdispersion and serial dependence [2].

Key words and phrases. Hidden Markov Model; Overdispersion; Serially-dependent; Simulation.

In this paper, we test the performance of the HMM for two sets of overdispersed data (dependent and non-dependent) using a simulation study. For this, we apply P-HMM to a series of monthly Hepatitis-B prevalence counts in the state of Kerala, South India, between 2006 and 2017. We estimate the P-HMM parameters and then decode the most likely sequence of hidden states.

2. Hidden Markov Models

An HMM is a two-phased stochastic process $Z_t = (H_t, O_t)$, the time variable $t \in N$, where the first stochastic process H_t is a finite set of states with a set of probabilities called transition probabilities and this state process is hidden refered to as parameter process. This hidden variable H_t satisfies Markov property, or it form a Markov chain (MC). The second stochastic process O_t is the set of observations which can be generated from any of these states. The distribution of O_t depends only on the current state H_t and not on previous states or observations. This process is called state-dependent process. Let $O_t = \{O_1, O_2, \cdots, O_T\}$ is the sequence of observed random variables and another set of discrete random variables $H_t = \{H_1, H_2, \cdots, H_T\}$ is the hidden or unobserved. Consider the realizations $o^{(t)} = (o_1, o_2, \cdots, o_t)$ which is generated from the unobserved or hidden variables $h^{(t)} = (h_1, h_2, \cdots, h_t)$. Here $\{h_k\}$ form a discrete MC. Then the probability of observed h from state i at time t can be expressed as

$$P(H_t = i | O^{(t)} = o^{(t)}). \tag{2.1}$$

These conditional probabilities are called state probabilities and here $i = 1, 2, \cdots, m$ is the number of hidden states of the MC. The probability of transitions between the states can be expressed as

$$a_{ij} = P(H_t = j | H_1 = r, H_2 = s, ..., H_{(t-1)} = i) = P(H_t = j | H_{(t-1)} = i). \tag{2.2}$$

Theses probabilities $a'_{ij}s$ are called transition probabilities, which can also be represented in the form of a matrix $\mathbf{A} = (a_{ij})_{m \times m}$, known as the transition probability matrix (TPM) and it satisfies the following conditions: $a_{ij} \geq 0, i, j \geq 1$ and $\sum_{j=1}^{m} a_{ij} = 1, i = 1, 2, \cdots, m$.

Another important term in HMM is state dependent probabilities which can be defined as $P = (p_i(o))$, where

$$p_i(o) = P(o_t = o | h_t = i), \ i = 1, 2 \cdots, m. \tag{2.3}$$

This is the probability mass function (pmf) of O_t when the process is in state i. This expression can be conveniently represent in a matrix form.
Define the initial distribution $\pi = \pi_i$, where $\pi_i = P(H_t = i)$ as the probability that the state i at time t(a row vector with nonnegative elements) if

$$\pi A = \pi \ and \ \pi \mathbf{1}' = 1.$$

For discrete-valued observations O_t, for $t = 1, 2, 3, \cdots$, we have

$$P(O_t = o) = \sum_{i=1}^{m} P(H_t = i) P(O_t = o | H_t = i) = \sum_{i=1}^{m} \pi_i(t) p_i(o).$$

This expression can conveniently be rewritten in matrix form as:

$$= \pi \mathbf{P}(t) \mathbf{1}'.$$

Now, the joint distribution of HMM can be written as

$$p(o_1, o_2, ..., o_T; h_1, h_2, ..., h_T) = p(h_1)p(o_1|h_1)\prod_{t=2}^{T} p(h_t|h_{(t-1)})p(o_t|h_t).$$

Summing over $h_1, h_2, \cdots, h_T$ we have the marginal pmf of the observation sequence (2.4)

$$P(O_1, O_2, \cdots, O_T) = \sum_{(h_1)}\sum_{(h_2)}\cdots\sum_{(h_T)}\{\pi_{(h_1)}p(o_1|h_1)\prod_{(t=2)}^{T} p(h_t|h_{(t-1)})p(o_t|h_t)\}.$$

The likelihood function in matrix form is hence L_T given by

$$L_T = \pi P(o_1)AP(o_2)AP(o_3)\cdots AP(o_T)1'. \tag{2.5}$$

The three parameters (components) for HMM are state transition probabilities, state dependent probabilities and initial state distribution. Hence, an HMM can be represented as $\Lambda = (A, P, \Pi)$. Clearly, HMM is well defined by states, state probabilities, transition probabilities, state dependent probabilities and initial probabilities [2].

There are three main problems involved in an HMM analysis. Firstly, evaluation problem, for given observation sequence and the model, to compute $P(O|\Lambda)$, the probability of the observation sequence, which we can solve using forward-backward algorithm [3]. Secondly, the learning problem, where the likelihood L_T estimation will be performed, this can be done by using EM algorithm [4]. And thirdly, decoding problem, in which we try to uncover the hidden part of the model, ie, the most likely state sequence, using Viterbi algorithm [5].

3. Simulation Framework

We demonstrate the performance of the HMM for overdispersed and dependent count data through a simulation study, which is one of the objectives of this work. To achieve this goal, we generate an overdispersed and independent and overdispersed and serially dependent data from a bimodel Poisson distribution with mean values 5 and 20. Further, P-HMM is applied to the both the generated data to evaluate the model fit and estimation efficiency. The simulation is conducted with R software. Figure 1 & 2 shows the auto correlation function (ACF) of both the sets of generated data. Figure 1 clearly indicates that the data are serially dependent.

For both the settings, 1,000 runs are conducted. In each run, data are estimated via the EM algorithm method. The simulations are repeated on different sample sizes ranging from 300 to 900. Table 1 and 2 show the simulation results for two types of overdispersed counts with the sample sizes 300 to 900. The tables include the Mean Squared Errors (MSE) and the biases of the parameter estimates. The formulas of bias and MSE are as follows

$$bias(\hat{\theta}) = \bar{\theta} - \theta,$$

$$MSE(\hat{\theta}) = \frac{1}{N}\sum_{i=1}^{N}(\hat{\theta}_i - \theta)^2,$$

where θ is the true parameter value, $\hat{\theta}$ is the estimate of θ for the i^{th} simulated data, $\bar{\theta} = \frac{1}{N}\sum_{i=1}^{N}\hat{\theta}_i$, and N = 1000 is the number of replicates. We can see that table 1 have the lowest bias and MSE values in almost all cases.

TABLE 1. MSE (bias) values of estimates for overdispersed and serially dependent count data.

N	300	500	700	900
a_{11}	0.0029(0.0466)	0.0027(0.0485)	0.0027(0.0492)	0.0026 (0.0486)
a_{21}	0.0028(-0.0461)	0.0027(-0.0479)	0.0026(-0.0481)	0.0027(-0.0497)
λ_1	0.0363(-0.0047)	0.0223(-0.0030)	0.0156(-0.0095)	0.0122(0.0044)
λ_2	0.1468(-0.0055)	0.0813(-0.0043)	0.0569(0.0041)	0.0458(-0.0064)

TABLE 2. MSE (bias) values of estimates for overdispersed count data.

N	300	500	700	900
a_{11}	0.2061(-0.4517)	0.2043(-0.4505)	0.2051(-0.4518)	0.2039(-0.4508)
a_{21}	0.0651(0.2518)	0.0633(0.2499)	0.0637(0.2511)	0.0632(0.2505)
λ_1	0.0529(0.0064)	0.0307(-0.0069)	0.0219(-0.0013)	0.0172(0.0033)
λ_2	0.1191(-0.0234)	0.0719(-0.0024)	0.0514(0.0097)	0.3966(-0.0081)

We known that count data are basically modelled by Poisson. But in many practical situations overdispersion exists and hence, we go for Negative Binomial

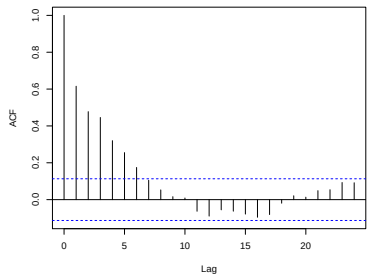

FIGURE 1. Overdispersed and serially dependent series: sample autocorrelation function.

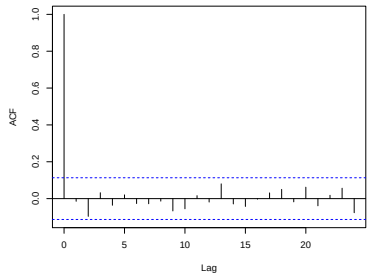

FIGURE 2. Overdispersed and independent series: sample autocorrelation function.

TABLE 3. Model selection statistics

model	-LogL	AIC	BIC
Poisson	1477.34(1390.39)	2956.68(2782.78)	2960.39(2786.49)
NegBin	1020.17(1047.91)	2044.35(2099.81)	2051.76(2107.22)
P-Mixture	948.57(974.51)	1903.14(1955.01)	1914.25(1966.12)
P-HMM	829.13(973.50)	1668.27(1957.00)	1686.79(1975.52)

(NegBin) or Poisson mixture (P-Mixture) models for modelling such data. We compare P-HMM with Poisson, NegBin and P-Mixture models and the result obtained is shown in table 3. It shows the values of model selection statistics of overdispersed and serially dependent data and those of overdispersed and independent data (in parenthesis). Summing up, we can say that HMM is the best model for fitting, when the data is overdispersed and serially dependent. When the observations are not correlated, mixture models are also preferable for modelling.

4. Real data analysis using P-HMM

A series of monthly Hepatitis-B incidence counts in the state of Kerala in South India for the 2006-2017 period is considered for the analysis. A time series plot of this data is given in figure-3. It has as many as 144 time points. A total of 10743 cases were studied with mean value 74.6 and variance 1370.8. In this case, the sample variance is greater than its mean, which shows that the data is clearly overdispersed. In this Hepatitis-B time series data the minimum occurrence of 8 incidences was recorded in *October* 2007 while the maximum occurrence of 222 was reported in *June* 2012.

Figure 4 shows the ACF of the Hepatitis-B incidence counts between 2006 and 2017. As evident in the figure, the observations are considerably correlated over the lags that gives direct serial dependence in the Hepatitis-B time series. The overdispersion and serial dependence of the incidence counts of Hepatitis-B time series have prompted us to build a P-HMM for analyzing this series.

Here the parameters are estimated from Hepatitis-B counts by applying the EM estimation method using R software. The parameter estimation of P-HMM (m = 2,3,4,5,6 & 7) is shown in Table 4. We assume that the underlying MC is stationary. Since $\widehat{\pi}$ is the stationary distribution of TPM $\widehat{A}$ in each case, its value need not

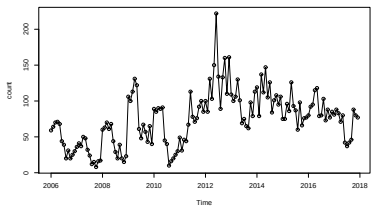

FIGURE 3. monthly Hepatitis-B cases in Kerala.

TABLE 4. P-HMMs fitted to the Hepatitis-B series

model	i	λ	A	Mean
m = 2	1	31.9	$\begin{bmatrix} 0.8434 & 0.1566 \\ 0.0709 & 0.9291 \end{bmatrix}$	93.37
	2	93.1		
m = 3	1	31.3	$\begin{bmatrix} 0.8547 & 0.1239 & 0.0214 \\ 0.0946 & 0.7728 & 0.1327 \\ 0.0000 & 0.2930 & 0.7070 \end{bmatrix}$	78.59
	2	78.6		
	3	122.8		
m = 4	1	31.2	$\begin{bmatrix} 0.8576 & 0.1183 & 0.0241 & 0.0000 \\ 0.1050 & 0.7114 & 0.1836 & 0.0000 \\ 0.0000 & 0.3223 & 0.5679 & 0.1098 \\ 0.0000 & 0.0000 & 0.7830 & 0.2170 \end{bmatrix}$	76.09
	2	76.1		
	3	110.4		
	4	167.0		
m = 5	1	19.4	$\begin{bmatrix} 0.7270 & 0.1809 & 0.0401 & 0.0520 & 0.0000 \\ 0.2034 & 0.6014 & 0.1952 & 0.0000 & 0.0000 \\ 0.0000 & 0.1167 & 0.7027 & 0.1806 & 0.0000 \\ 0.0000 & 0.0000 & 0.3306 & 0.5552 & 0.1142 \\ 0.0000 & 0.0000 & 0.0000 & 0.7893 & 0.2107 \end{bmatrix}$	77.31
	2	41.7		
	3	77.3		
	4	111.4		
	5	167.5		
m = 6	1	19.3	$\begin{bmatrix} 0.7220 & 0.1780 & 0.0473 & 0.0527 & 0.0000 & 0.0000 \\ 0.2071 & 0.5995 & 0.1934 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.1433 & 0.6781 & 0.1786 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.2557 & 0.4502 & 0.2941 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.8424 & 0.0803 & 0.0773 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 \end{bmatrix}$	74.48
	2	41.2		
	3	74.5		
	4	100.2		
	5	136.9		
	6	221.9		
m = 7	1	14.4	$\begin{bmatrix} 0.6755 & 0.2082 & 0.0000 & 0.1163 & 0.0000 & 0.0000 & 0.0000 \\ 0.0966 & 0.5997 & 0.2295 & 0.0000 & 0.0742 & 0.0000 & 0.0000 \\ 0.0647 & 0.1605 & 0.5613 & 0.2135 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.1432 & 0.6784 & 0.1784 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.2557 & 0.4498 & 0.2945 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.8424 & 0.0803 & 0.0773 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 \end{bmatrix}$	74.54
	2	25.1		
	3	42.6		
	4	74.5		
	5	100.3		
	6	136.9		
	7	221.9		

been estimated. So $\widehat{\pi}$ automatically approaches a unit vector.

Table 5 shows the AIC and BIC values for each model. In this table k represents the number of parameters of each model and is given by $k = m^2$. Of these, the $6-state$ model is chosen by AIC and $5-state$ by BIC. The model is considered to

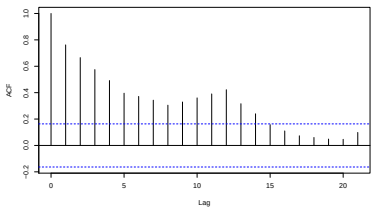

FIGURE 4. The autocorrelation function of the Hepatitis-B series.

TABLE 5. Model selection statistics for P-HMM

model	k	-LogL	AIC	BIC
m = 1	1	1818.11	3638.22	3641.19
m = 2	4	940.62	1889.25	1901.13
m = 3	9	759.63	1537.25	1563.98
m = 4	16	721.44	1474.88	1522.40
m = 5	25	652.62	1355.23	**1429. 48**
m = 6	36	633.62	1339.24	1446.15
m = 7	49	625.48	1348.96	1494.48

TABLE 6. The most likely sequence of hidden states of $5-state$ P-HMM

3	3	3	3	3	2	2	1	1	1	1	2	2	2	2	2	2	2	1	1	1	1	1	1	3	3	3
3	3	3	3	3	2	2	1	1	1	1	2	2	2	2	2	2	2	1	1	1	1	1	1	3	3	3
3	3	2	1	1	2	1	1	1	4	4	4	4	4	3	2	3	2	2	3	2	3	3	3	3	3	2
2	1	1	1	1	2	2	2	2	2	3	4	3	3	3	3	4	3	4	3	4	4	5	5	4	4	4
5	4	5	4	4	4	4	4	3	3	3	3	3	3	4	4	3	4	4	5	4	4	3	4	4	4	4
3	3	3	3	4	3	3	3	3	3	3	3	3	4	4	4	4	3	3	4	3	3	3	3	3	3	3
3	3	2	2	2	2	3	3	3																		

be the most apt one, which can be identified using BIC. So the $5-state$ model was used for our analysis and the details of the estimated values are shown in Table 4. Even the $6-state$ model can be used as a simple alternative for the analysis. Clearly, P-HMM with $5-state$ model is better fit for our Hepatitis-B prevalence count. We can call these $5-states$ as, state 1 - 'large number of zero counts', state 2 - 'mild occurrence', state 3 - 'moderate occurrence', state 4 - 'severe occurrence' and state 5 - 'inflation of outliers'. After determining the classification of the states of the P-HMM, the estimated parameters of $5-state$ model are processed by the Viterbi algorithm (as shown in Table 3) to get the sequence of hidden states. This gave as the best sequence of the hidden states which generates our observations.

Though Hepatitis B cases are a rarity in Kerala, a sudden outbreak of the disease was reported during the 2006-17 period. In such a scenario, HMM is the best method for modelling data. We analysed these observations using P-HMM and fitted them to the $5-state$ model. The parameters of this model is estimated using EM algorithm, as shown in table 4. The transition probability matrix gives the probability of the transition from one state to another at any time point. We can notice very strong diagonal elements in the transition probability matrix of the $5-state$ P-HMM. This means that if the data is generated from state i in this month, there is a strong possibility that the data for the next month will be generated from state i itself. This is evident from the state sequence we derived using Viterbi algorithm.

5. Discussion

As per the WHO guidelines, India is in the intermediate prevalence zone of Hepatitis-B infection. However, recent studies point to the fact that Hepatitis-B

virus prevalence in India is $< 2\%$, placing it in the low prevalence zone. Since Hepatitis-B virus is a transfusion transmissible infection, various studies were conducted on the prevalence of Hepatitis-B among blood donors [6].
The retrospective data over a span of 12 years from South India is considered in this study. In the case of data from a heterogeneous population, there is a possibility of clusters or mixtures in the occurrence of diseases. When there are excess zeros in the count time series data, zero inflated models may be an appropriate modelling method. One way of modelling the mixture count is HMM modelling. However this may change based on the scenario. In this study, the prevalence occurrences of diseases were modelled using the HMM because of the dependence nature of the disease.
AIC and BIC statistics were used for the model comparison. The models mentioned in this work can be viewed in the form of independent and dependent mixture models. The mixture indicated the disease occurrence level. The mixture occurrence in the data can be because of the seasonal or endemic pattern of the disease occurrence.
Researchers used various models to explain the pattern of Hepatitis-B virus disease occurrence. There are various infectious disease time series studies with ZIP and ZINB models. Held et al. 2005 ([7]) also used particular versions of count regression model with auto regressive term to incorporate the time series nature of the data. Researchers used different forms HMMs for modelling infectous diseases([8], [9], [10], [11]). When it comes to analysis of infectious diseases counts, ARIMA ([1]) and GLM ([12], [13]) methods are the ones used generally. P-HMM is one of the popular model for diseases survelliance as it is common to measure the infectious diseases as counts. Sarvi et al. 2017 ([14]) made inferences based on the air pollution data using P-HMM.
HMMs are popular became a traditional modelling nowadays. The models can answer many problems using the MC concepts, which cannot be obtained using the usual time series models such as ARIMA or SARIMA models. One of the most infectious and dangerous diseases, Hepatitis-B had been a major health concern across the globe. However, with the advent of medical tech and awareness, the prevalence level of the infection has declined over the years. This is evident from the decrease in death cases due to the infection in the state of Kerala in South India. Launched in 1982, the Hepatitis-B vaccine is 95% effective in preventing infection.

6. Conclusion

In this study, we conducted a simulation study for testing the HMM's performance for modelling overdispersed and serially dependent nature of observations. We can, beyond any doubt, say that HMM is the best model for fitting such data. Taking into account this particular scenario, the HMM was used for analysing the Hepatitis-B prevalence data, which are overdispersed and serially dependent. Here, P-HMM was applied for studying discrete time series count data, with which the hidden state of the observations could be easily traced out. Since the initial distribution was stationary and EM was used for the estimation of the parameters, the maximum value of the likelihood decreased. For all the models, the estimates of the transition probabilities and the state-dependent means were not the same, but close. Also, the estimated value of π reaches unit vector. Applying the AIC and

BIC methods, it was found that the 5-state model was the best fit for this data. Global decoding is used for finding out hidden state sequence in majority of the applications. Here, Viterbi algorithm is applied to determine the most likely state sequence from the 5-state model.

References

[1] W.D. Wei, J.J. Jiang, H. Liang, L. Gao, B.Y. Liang, J.G. Huang and et al., Application of a combined model with autoregressive integrated moving average (ARIMA) and generalized regression neural network (GRNN) in forecasting hepatitis incidence in Heng County, China, Plos One. 11 (1967), e156768.

[2] W. Zucchini and I.L. MacDonald, Hidden Markov Models for Time Series: An Introduction Using R, CRC Press, 2009.

[3] L.E. Baum, An Inequality and Associated Maximization Technique in Statistical Estimation for Probabilistic Functions of Markov Processes, Inequalities. 3 (1972), 1–8.

[4] A. P. Dempster, N. M. Laird and D. B. Rubin, Maximum Likelihood from Incomplete Data via the EM Algorithm, Journal of the Royal Statistical Society, Ser. B. 39 (1977), 1–38.

[5] A. J. Viterbi, Error Bounds for Convolutional Codes an and Asymptotically Optimum Decoding Algorithm, IEEE Trans.Info.Th.IT. 13 (1967), 260–269.

[6] P. Jadeja, A. Kaur, H. Shekha, Trend in seroprevalence of Hepatitis B Virus infection among blood donoers at a tertiary care center of Rajasthan, India, Natl J Med Res. 4 (2014), 205–207.

[7] L. Held, M. Hohle and M. Hofmann, A statistical framework for the analysis of multivariate infectious disease surveillance counts, Statistical Modelling. 5 (2005), 187–199.

[8] A. Rafei, E. Pasha and R. Orak, Tuberculosis surveillance using a hidden Markov model, Iran J Public Health. 41 (2012), 87–96.

[9] T. Sebastian, V. Jeyaseelan, L. Jeyaseelan, S. Anandan, S. George and S. I. Bangdiwala, Decoding and modelling of time series count data using Poisson hidden Markov model and Markov ordinal logistic regression models, Statistical Methods in Medical Research. 28 (2019), 1552–1563.

[10] B. Cooper and M. Lipsitch, The analysis of hospital infection data using hidden Markov models, Biostatistics. 5 (2004), 223–237.

[11] L.Y. Strat and F. Carrat, Monitoring epidemiologic surveillance data using hidden Markov models, Stat. Med. 18 (1999), 3463–3478.

[12] A. Agresti, An Introduction to Categorical Data Analysis, John Wiley & Son, Inc. 2007.

[13] S. Chandy, K. Ramanathan, A. Manoharan, and et al., Assessing effect of climate on the incidence of dengue in Tamil nadu, Indian J. Med Microbiol. 31 (2013), 283–286.

[14] F. Sarvi, A. Nadali, M. Khodadost, M. K. Moghaddam and M. Sadeghifar, Application of Poisson Hidden Markov Model to Predict Number of PM 2.5 Exceedance Days in Tehran During 2016-2017, Avicenna J Environ Health Eng. 4 (2017), e.58031.

[1]Department of Statistics,St.Thomas College,Pala, Kerala, India

[2]Department of Biostatistics, Christian Medical College, Vellore, India

*Corresponding author: ambilystat06@gmail.com