Predicted Cost Model for Integrated Healthcare Systems Using Markov Process

Sangjin Jeong[1,2], Chan-Hyun Youn[3], and Yong-Woon Kim[2]

[1] Dept. of Information and Communications Engineering, KAIST, Daejeon, Korea
[2] Protocol Engineering Center, ETRI, Daejeon, Korea
{sjjeong,qkim}@etri.re.kr
[3] Dept. of Electrical Engineering, KAIST, Daejeon, Korea
chyoun@kaist.ac.kr

Abstract. Predicting cost in the health care environment is a challenging issue for healthcare professionals. Chronic disease is a long term disease that requires life time care. Physicians need to keep tracking patients' status over time including routine medical examinations. In this paper, we investigate the healthcare service scenario for chronic disease care. Then, we propose a predicted cost model for the investigated healthcare service. The cost model is based on the prediction of utilization and attendant costs through the development of a stochastic model, specifically a first-order Markov chain, can be adapted to specific diseases and/or events. The proposed model is at the initial stage and may require testing using a large administrative database of patients and hospital operational costs.

Keywords: healthcare, chronic disease, metabolic syndrome, service level agreement.

1 Introduction

Predicting cost in the health care environment is a challenging dilemma for medical professionals. The importance of a viable cost model incorporating outcomes measurement and payment schemes is of interest. Healthcare administrators want to assure that the delivery of services is appropriate as identified by government guidelines, rules and regulations. A critical starting point is to provide the framework necessary to provide a cost model that considers the general factors of healthcare encounters, patient diagnosis, treatment and the related costs that can be used to describe this complex problem. The stochastic nature of disease treatment can lead to substantial variation in experience between and among classes of enrollees, their diseases, and treatment utilization patterns. The most common approach to analyzing cost of healthcare service is the traditional method of summing the number of events occurring in the system over a period of time and calculating the mean and a standard deviation of cost. However, due to the emergence of new healthcare service models such as integrated healthcare system, there is an increasing demand for more sophisticated models to predict cost in the healthcare environments. A wide variety of

Y.-S. Jeong et al. (eds.), *Ubiquitous Information Technologies and Applications*,
Lecture Notes in Electrical Engineering 280,
DOI: 10.1007/978-3-642-41671-2_24,

conceptual and statistical models exist, both deterministic and stochastic, to measure utilization in health research. The typical types of deterministic models are traditional summing of events-based model, and decision tree-based decision analysis model. However, it is known that the limitation of the traditional models is its inability to account for non-symmetric aspect of cost data and the lack of consideration of the utilization patterns of the population. Decision tree-based models can be effective models in economic and policy analyses, because they can provide information to patients and practitioners about risk and cost. The difficulties with this model arise when timing becomes a concern. This problem becomes apparent when the time interval is several years or there are repeated events in a shorter time interval [7].

To resolve the limitations above, the Markov process has been widely used for modeling of epidemic progression of various diseases such as influenza, tuberculosis and HIV, and the pathways for subjects to predict utilization or possible pathways through the healthcare system [1-5]. Among them, to predict utilization in various healthcare plans and healthcare systems, Kapadia et al. studied 305 patients at a rehabilitation hospital over a six months period [5]. The authors developed a hospital service charging model by measuring the utilization patterns of the patients. Also, Beland studied ambulatory care in Canada. The author used the physician claims from clinic visits, hospitalization and emergency room visits and adopted a Markov chain to predict utilization to show the differences between population demographics such as age and gender. The author showed that the corresponding changes in the traditional model of counting visits to physicians can be modeled using Markov process. The results of the above literature review indicate that the Markov process is appropriate to estimate the utilization of a population of patients or enrollees. The Markov process can illustrate the difference in the treatment utilization patterns due to predictor variables such as gender and age [4]. Also, Leviton et al. examined the application of a generalized Markov process seems appropriate to predict utilization for patients with chronic or acute diseases [6]. In accordance with the previous literature, in this paper, we develop a Markov process-based cost prediction model for integrated healthcare system in home-hospital environment.

2 Markov Process Modeling for Integrated Healthcare Service

This section presents a predicted cost model for service scenario of home-hospotal environment integrated healthcare systems.Table 1 shows the states of integrated healthcare system in home-hospital environment used in this subchapter to develop the Markov process. The states represent typical disease management scenario of integrated healthcare system [7].

The Markov process is being employed as the first component of this model to predict utilization for this healthcare problem. The transition or change in utilization from state i to state j is influenced by the prior state is denoted as:

$$P_{ij} = \Pr[X_{n+1} = j \mid X_n = i], \tag{1}$$

where P_{ij} is the probability of going from state i to state j in one step or one increment in the time unit.

Table 1. States of integrated healthcare system in home-hospital environment for the Markov process

State 0 (m_0)	No identified chronic disease
State 1 (m_1)	Self-management at home
State 2 (m_2)	Consult an informant about abrupt change on at-home medical examination results
State 3 (m_3)	Visit hospital (periodic visit or abrupt disease state change)
State 4 (m_4)	Diagnosis by physician
State 5 (m_5)	Laboratory tests
State 6 (m_6)	Discharged from the hospital
State 7 (m_7)	Complete cure

Due to the patient's outpatient flow is managed by predetermined patient management care plan of hospital, it is not necessary to consider all the previous transitions, when determining next transition. All the transitions from the resident states require preset exit criteria and only the previous state influences the opportunity for transition to an alternate state. This means that being in a resident utilization state two transitions earlier is irrelevant to the state you are in presently. All the information that is needed is the previous transition state. This is using the memory less or Markovian property. The resident states of transition based on utilization for the Markov process will be defined as follows: The states defined in Table 1 lead to an eight states Markov chain with an absorbing barrier (State 7) and the resulting one step transition probability matrix is

$$\mathrm{P}_{7\times 7} = (P_{ij}), \text{ where } i = 0,1,2,\ldots,7 \text{ and } j = 0,1,2,\ldots,7 \tag{2}$$

We specify the time interval unit for the probability transition matrix for utilization to be one unit time T.

Consider the finite number of possible transitions for individuals in the model. Denote by the element

$$P_{00} = \Pr[X_n = 0 \mid X_{n-1} = 0], \text{ for any } n. \tag{3}$$

where the probability that an individual starting in m_0 stayed in m_0 after one time period.

Similarly, the probabilities for staying in the same state after one transition are denoted as follows:

$$P_{00}, P_{11}, P_{22}, P_{33}, P_{44}, P_{55}, P_{66}, P_{77} \tag{4}$$

$$\text{Let } f_{ii}^{n} = \Pr[X_n = i, X_v \neq i, v = 1,2,\ldots,n-1 \mid X_0 = i] \tag{5}$$

be the probability that, starting from state i, the first return to state i occurs at the nth transition [8]. There are 64 possible transitions for an individual in this model. State m_7 (complete cure) is an absorbing barrier or state, defined as a state that once entered cannot be exited [8]. This means that the probabilities of transition from m_7 to states m_0 through m_6 are zero and the probability of starting in state seven and staying in state seven is one. Hence in the probability transition matrix Eq. (2) becomes

$$P_{70} = P_{71} = P_{72} = P_{73} = P_{74} = P_{75} = P_{76} = 0 \;\&\; P_{77} = 1 \tag{6}$$

The matrix can be partitioned into 4 subsets. The set of transition probabilities, $\{P_{77}\}$, has the following two properties:

1. It has a period of one
2. Since $f_{77}^{n} = 1$, it is positive recurrent.

Combining the previous two properties leads to the conclusion that $\{P_{77}\}$ is an ergodic set. This set will be represented by the sub-matrix, $\mathrm{E}_{1\times 1}$ and defined as follows:

$$\mathrm{E}_{1\times 1} = \{P_{77}\} \tag{7}$$

The next submatrix of the partitioned matrix to be considered is the vector of zeroes,

$$\mathrm{O}_{1\times 7} = \{P_{7\mathrm{i}} \mid i = 0, 1, 2, 3, 4, 5, 6\}. \tag{8}$$

Once in the absorbing state the individual cannot leave the state; hence all the cell entries axe zero. The third partitioned sub-matrix to be defined will be the transient states,

$$\mathrm{M}_{7\times 7} = \{P_{\mathrm{ij}} \mid i = 0, 1, 2, 3, 4, 5, 6 \;\&\; j = 0, 1, 2, 3, 4, 5, 6\}. \tag{9}$$

The sub-matrix M includes all the transient states of the Markov chain. The probability of the first return, Eq. (5), for these states (m_0, m_1, m_2, m_3, m_4, m_5, m_6) is less than one. The final sub-matrix to consider is the transition from a transient state to the absorbing state. This will be defined as

$$\mathrm{L}_{7\times 1} = \{P_{ij} \mid i = 0, 1, 2, 3, 4, 5, 6\} \tag{10}$$

An alternative form of the probability transition matrix can now be illustrated with dimensions of partitioned matrices:

$\mathrm{E}_{1\times 1}, \mathrm{O}_{1\times 7}, \mathrm{L}_{7\times 1}, \mathrm{M}_{7\times 7}$. It should be noted that the matrix $\mathrm{E}_{1\times 1}$ is equivalent to the identity matrix, $\mathrm{I}_{1\times 1}$. Replacing $\mathrm{E}_{1\times 1}$ with $\mathrm{I}_{1\times 1}$ in the matrix results in the following

$$\mathrm{P} = \begin{pmatrix} \mathrm{I}_{1\times 1} & \mathrm{O}_{1\times 7} \\ \mathrm{L}_{7\times 1} & \mathrm{M}_{7\times 7} \end{pmatrix} \tag{11}$$

3 Predicted Cost Model for Integrated Healthcare Systems Service Scenario

Determining the mean time or number cycles an individual occupies in a resident state requires some knowledge of linear algebra and the development of the fundamental matrix for Markov chain with an absorbing state. Kemeny and Snell [9] developed methodology for finding the mean time in each resident state before transition into the absorbing state. They proposed the following.

Let $\mathrm{B}_{n\times n}^{m}$ be a square matrix raised to the power *m*. If $B^m \to 0$ as $m \to \infty$, then $(\mathrm{I}-\mathrm{B})$ has an inverse, and

$$(\mathrm{I}-\mathrm{B})^{-1} = \mathrm{I}+\mathrm{B}+\mathrm{B}^2+\ldots = \sum_{i=0}^{\infty}\mathrm{B}^i \tag{12}$$

For any Markov chain with an ergodic set, let the matrix M correspond to the set of transient states, as in Eq. (11). Then $(\mathrm{I}-\mathrm{M})$ has an inverse, and

$$(\mathrm{I}-\mathrm{M})^{-1} = \mathrm{I}+\mathrm{M}+\mathrm{M}^2+\ldots = \sum_{i=0}^{\infty}\mathrm{M}^i \tag{13}$$

Substituting the matrix M from Eq. (11) into Eq. (12) proves Eq. (13).

Let

$$\mathrm{N} = (\mathrm{I}-\mathrm{M})^{-1} \tag{14}$$

be the fundamental matrix for a Markov chain with an ergodic state [9]. The next consideration is the number of times for an individual that a transient state is occupied. Define η_{lj} to be the function assigning the total number of times that the process is in state m_j after starting from state m_l (restricting the choices to transient states, $\{m_j \mid j = 0,1,2,3,4,5,6\}$. This quantity will be will be expressed as the sum of indicator variables

$$\mu_{lj}^{k} = \begin{cases} 0, \text{if the process is in state } m_j \text{ after } k \text{ steps} \\ 1, \text{otherwise} \end{cases} \tag{15}$$

Determining the expectation of the number of cycles an individual stays in a resident transition state, conditional on having just entered the system, follows [9]. The mean number of days spent in m_j after starting in state m_l is $\mathrm{N}_{7\times 7} = E[\eta_{lj}]$ as can be seen from the following argument. It should be observed that $\eta_{lj} = \sum_{k=0}^{\infty}\mu_{lj}^{k}$. Hence,

$E[\eta_{lj}] = E\left[\sum_{k=0}^{\infty} \mu_{lj}^{k}\right]_{7\times 7}$. Note that the μ_{lj}^{k} the l, j element of M^{k}. Here η is the matrix whose l, j element is M^{k}. Then

$$E[\eta] = \sum_{k=0}^{\infty} E[\mu_{lj}^{k}] = \sum_{k=0}^{\infty} \mathrm{M}^{k} = \mathrm{N} \tag{16}$$

Denote the expected numbers of days in seven transient states by T', taken from the proper row of *N*.

The notation for the cost function is Eq. (17). Define the fixed cost to be a column vector, where each element of this 8×1 matrix is the averaged costs per utilization state of the system. It should be noted that the elements for the states "no use of services" and "Complete cure" have no allowable costs associated with them. The cost function can be represented as Eq. (17)

$$C'_{8\times 1} = \{0, c_1, c_2, c_3, c_4, c_5, c_6, 0\} \tag{17}$$

The model is defined by multiplying T' and Eq. (17) with the result

$$\mathrm{F}(x_i) = \mathrm{T}'_{8\times 1}(x_i)\mathrm{C}_{1\times 8}, \tag{18}$$

where x_i is the conditions of interest (gender, age, and diagnosis).

The value of the function F is the predicted cost for an individual. The vector of utilization, T', has a dimension of 1×8. The vector of cost, C, has a dimension of 8×1. Taking their product generates a scalar value, $\mathrm{F}_{1\times 1}$, which is the predicted cost given gender, age, and diagnosis.

4 Conclusion

Predicting cost in the health care environment is a challenging issue for healthcare system professionals. In this paper, we presented Markov process-based cost model to predict costs based upon healthcare service scenario of home-hospital integrated environments. The proposed cost model has following limitations. If the number of transitions is small in one or more resident states with the addition of one or more resident states becoming ergodic, then an unstable probability transition matrix is generated. The unstable matrix cannot provide appropriate estimates. The proposed model has not been tested on a large administrative database of claims. However, during the test phase of the proposed model, the following issues need to be considered. The first issue is that the restrictions due to the database with the lack of demographic identifiers. The absence of ethnicity and marital status may lead to questions about the changes in utilization patterns for these groups of enrollees. A second issue about evaluation is the choice of deleting the multiple events per day for an individual. This may cause the states remaining transient in the probability transition matrices. Therefore, the issues above need to be considered during further evaluation of the proposed model.

Acknowledgment. This research was supported by the ICT Standardization program of MKE (The Ministry of Knowledge Economy).

References

1. Aalen, O.O., Farewell, V.T., De Angelis, D., Day, N.E., Gill, O.N.: A Markov model for HIV disease progression including the effect of HIV diagnosis and treatment: Application to AIDS prediction in England and Wales. Statistics in Medicine 16, 2191–2210 (1997)
2. Auranen, K., Ranta, J., Takala, A.K., Arjas, E.: A statistical model of transmission of Hib bacteria in a family. Statistics in Medicine 15, 2235–2252 (1996)
3. Beck, J.R., Paulker, S.G.: The Markov process in medical prognosis. Medical Decision-Making 3(4), 419–458 (1983)
4. Beland, F.: Conceptualizing the utilization of ambulatory medical care as a process. Medical Care 26(2), 115–123 (1988)
5. Kapadia, A.S., Vineberg, S.E., Rossi, C.D.: Predicting course of treatment in a rehabilitation hospital: A Markovian model. Computers & Operation Research 12(5), 459–469 (1985)
6. Leviton, A., Schulman, J., Kammerman, L., Porter, D., Slack, W., Graham, J.R.: A probability model of headache recurrence. Journal of Chronic Disease 33, 407–412 (1980)
7. McGhee, C.R., Glasser, J.H., Chan, W., Pomeroy, N., Chan, W.: Forecasting Health Care Expenditures and Utilization Based on a Markov Process and a Deterministic Cost Function in Managed Care Settings. Lecture Notes-Monograph Series, vol. 43, pp. 229–238. Statistical Essays in Honor of Jack Hall, Crossing Boundaries (2003)
8. Ross, S.M.: Introduction to Probability Models, 3rd edn. Academic Press, San Diego (1997)
9. Kemeny, J.G., Snell, J.L.: Finite Markov Chains. D. Van Nostrand Company, Inc., Princeton (1960)