International Journal of Business, Human and Social Sciences
ISSN: 2517-9411
Vol:6, No:10, 2012

A Novel Approach to Handle Uncertainty in Health System Variables for Hospital Admissions

Manisha Rathi, Thierry Chaussalet

***Abstract*—**Hospital staff and managers are under pressure and concerned for effective use and management of scarce resources. The hospital admissions require many decisions that have complex and uncertain consequences for hospital resource utilization and patient flow. It is challenging to predict risk of admissions and length of stay of a patient due to their vague nature. There is no method to capture the vague definition of admission of a patient. Also, current methods and tools used to predict patients at risk of admission fail to deal with uncertainty in unplanned admission, LOS, patients' characteristics.

The main objective of this paper is to deal with uncertainty in health system variables, and handles uncertain relationship among variables. An introduction of machine learning techniques along with statistical methods like Regression methods can be a proposed solution approach to handle uncertainty in health system variables. A model that adapts fuzzy methods to handle uncertain data and uncertain relationships can be an efficient solution to capture the vague definition of admission of a patient.

***Keywords*—**Admission, Fuzzy, Regression, Uncertainty

I. Introduction

THE number of hospital admissions in UK National Health Service (NHS) and in several other developed countries has been rising for many years [4]. Hospital admissions of patients tend to be vaguely defined due to unexpected variations in the data, patients' characteristics, demographic variables and indeed health system variables. Future admission of a patient is uncertain; therefore it is challenging to predict patients at high risk of admission. There are a number of modeling approaches and statistical methods to predict unplanned admissions and length of stay of a patient. There is uncertainty in unplanned admissions, LOS, patients' characteristics and most variables describing the health system. These models are useful in providing insights into the behavior of patients' admission. Unplanned admission of a patient can be at high risk or low risk [5]. These categories are vague or fuzzy. Risk factors (clinical, social factors, patients' characteristics and demographic characteristics) are associated with determining the risk of unplanned admission of a patient. However, these models fail to deal with uncertainty in these risk factors. In addition, uncertain relationship exists between data variables. Not all data can be captured by linear type of relationships. The type of relationships among variables can be linear or non-linear. Therefore, relationships among variables are uncertain and it becomes difficult to handle such relationships. Precise and accurate models are very important in handling vague nature of admission of a patient.

Manisha Rathi is with the , University of Westminster in School of Electronics and Computer Science as PhD student (email manisha.rathi@my.westminster.ac.uk).

Prof Thierry Chaussalet, is professor with University of Westminster in School of Electronics and Computer Science (e-mail: chaussalet@westminster.ac.uk).

Studies have found that admission of a patient is a dichotomous response variable. A problem may arise in prediction due to the vague nature of the response variable. The purpose of this paper is to develop a model to handle uncertainty in patients' characteristics and risk factors associated with unplanned hospital admissions.

II. Motivation of Paper

Patients at high risk of unplanned admission could be targeted for interventions designed for efficient hospital resource utilization [6]. A number of predictive tools for the prediction of patients who are at high risk of readmission have been developed [2] [5].

Most Predictive models have focused on regression techniques, although there is an emerging interest in artificial intelligence. There are limitations and drawbacks of the current predictive models. Problems may arise due to uncertainty in data variables, uncertain relationship among variables and vagueness in the admission of a patient.

The main challenges of this paper are:

- To handle uncertainty in unplanned admission and patients' characteristics.
- To handle uncertain relationships among variables.
- To capture the vague definition of unplanned admission of a patient.

III. Proposed Model

A model is developed to capture the vague definition of unplanned admission. This model also handles uncertainty in unplanned admission, patients' characteristics and health system variables. Admission of a patient can be at high, medium or low risk. These values high, medium or low are fuzzy and can be placed in a fuzzy set.

Admission of a patient is a fuzzy event because it can take the values other than 0 or 1 and unplanned admissions do not have clear cut boundaries. This model identifies patients' characteristics and risk factors associated with unplanned hospital admission. Due to uncertainty in available information, the future state (admission of a patient) is not clearly defined.

The model can be generated by adapting two methods: Fuzzy techniques and regression methods. This model is adapted when available data is uncertain and when explanatory variables are interacting in uncertain manners.

Adapting machine learning technique can be a proposed approach to handle uncertain relationship and unknown dependency between a given set of input variables and its response variables.

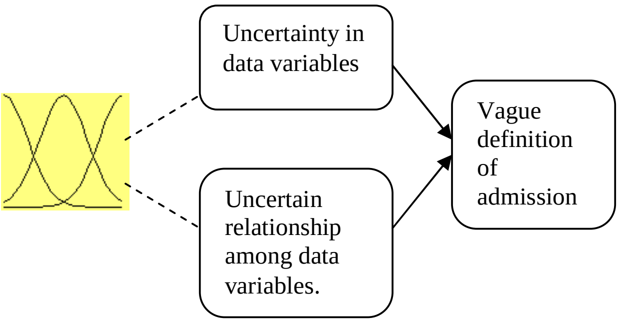

Fig. 1 Shows Proposed Approach for capturing vague definition of admission

IV. Statistical Method and Analysis

Statistical method and analysis were performed using SPSS. Regression Analysis can be used to assess the associations between admission method and independent variables including admission source and age. The test for normal distribution is conducted through Probability-Probability Plot (P-P). The plot is a graph of the empirical CDF values plotted against the theoretical CDF values. The plot will be approximately linear if the specified theoretical distribution is the correct model.

V. Data Collection

Information on risk factors associated with unplanned admission was obtained from HES datasets. For, the cases data were collected for the admission methods, admission source, Hospital provider spell no, Spell duration, start age of patient, end age of patients and primary diagnosis. Regression analysis can be run on the sample data set to analyze relationships between the independent and dependent variables, and understand the effect of predictor variables on the outcome.

VI. Correlation Regression

Regression analysis has been run on the sample data set. The table I show the correlation among the variables. The value close to 1 indicates that there is a strong relationship among the variables. The value closes to 0 shows weak relationship between variables. The positive value of Pearson's r means that if one variable increases in value then the other variable also increases in value. Similarly, the negative value of Pearson's r means that if one variable decreases in value than other variable also decreases in value. Sig (1-tailed) value is .206, the value is greater than 0.05. That means, increases or decreases in one variable do not significantly relate to increases or decreases in your second variable. The other value is 0.001; we can conclude that there is a statistically significant correlation between two variables.

That means, increases or decreases in one variable do significantly relate to increases or decreases in your second variable.

TABLE I
Shows the Correlation between Variables

		VAR00001	VAR00006	VAR00005	VAR00007
Pearson Correlation	VAR00001	1.000	.026	-.162	.098
	VAR00006	.026	1.000	.127	.910
	VAR00005	-.162	.127	1.000	-.216
	VAR00007	.098	.910	-.216	1.000
Sig. (1-tailed)	VAR00001	.	.206	.000	.001
	VAR00006	.206	.	.000	.000
	VAR00005	.000	.000	.	.000
	VAR00007	.001	.000	.000	.
N	VAR00001	997	997	997	997
	VAR00006	997	997	997	997
	VAR00005	997	997	997	997
	VAR00007	997	997	997	997

VII. Fuzzy Membership Function

This visualization is helpful to understand how the system is going to behave for the entire range of values in the input space. Upon opening the surface viewer, we are presented with the two-input one output case, we can see the entire mapping on one plot. The fuzzy membership function can be represented with the help of triangular or Gaussian type of functions. The functions can be represented by mathematical equations like

$$f(x;a,b,c) = \max(\min(\frac{x-a}{b-a}, \frac{c-x}{c-b}), 0) \quad (1)$$

The parameters a and c locate the "feet" of the triangle and the parameter b locates the peak.

The parameters for gaussmf represent the parameters σ and c listed in order in the vector [sig c].

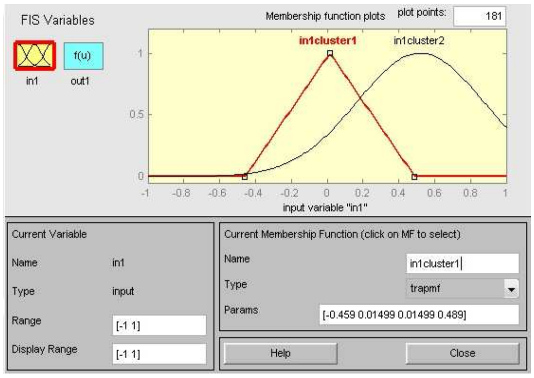

Fig. 2 shows triangular membership function plots (-.0.459, 0.459)

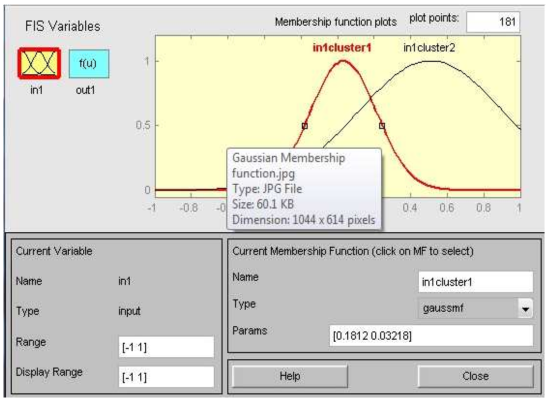

Fig. 3 shows the Gaussian membership function plot of variables in the range [-1,1] and fuzzy set (0.1812, 03218)

VIII. Conclusion and Future Work

It can be concluded that due to uncertainty in health system variables and uncertain relationship among variables, there is vagueness in the admission of a patient. A model that includes statistical techniques along with fuzzy techniques is a proposed approach to deal with uncertainty in data variables and uncertain relationship among variables.

It can be analyzed from P-Plot that not all data can be captured linear type of relationships and all data is not normally distributed. Also, it can be shown by regression analysis that strong and statistically significant relationship exists between some variables whereas there is a weak relationship between some variables and correlation between some variables is not statistically significant. Due to uncertainty in health system variables , the fuzziness in data variables occurs, which can be plotted with the help of fuzzy membership functions.

IX. Results

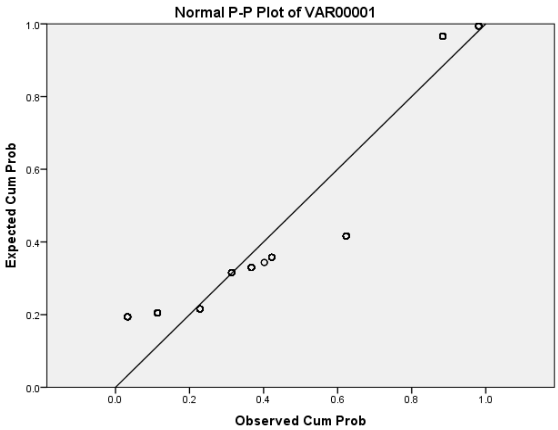

Fig. 4 Shows the P-Plot of VAR00001, data can be captured by linear relationships but not all data is normally distributed

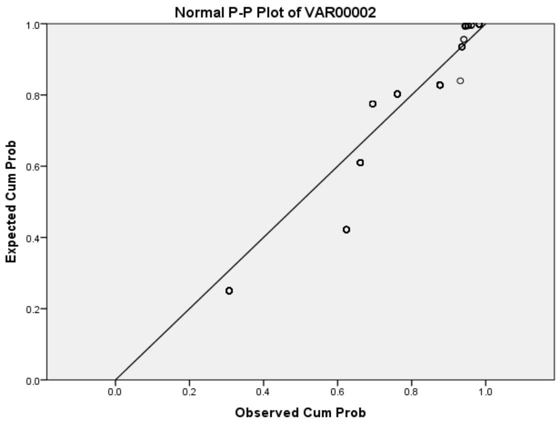

Fig. 5 Shows the P-Plot VAR00002, data can be captured by linear relationship but not all data is normally distributed

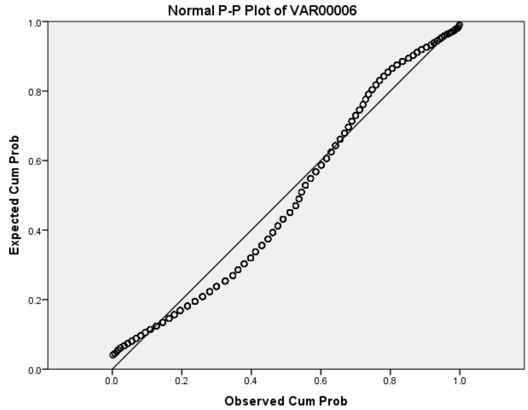

Fig. 6 shows the P-Plot of VAR00006, linear relationships do not exist; data cannot be captured by just correlation and Regression

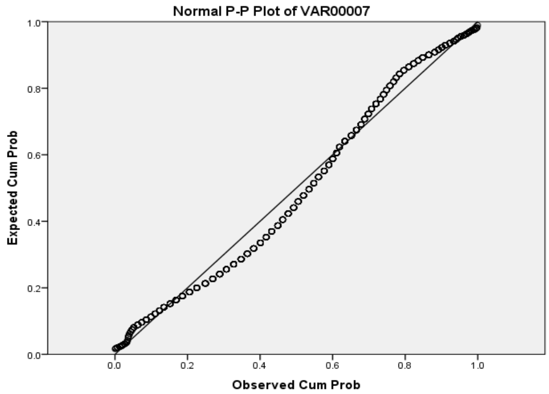

Fig. 7 shows the P-Plot of VAR00006, linear relationships do not exist; data cannot be captured by just correlation and Regression

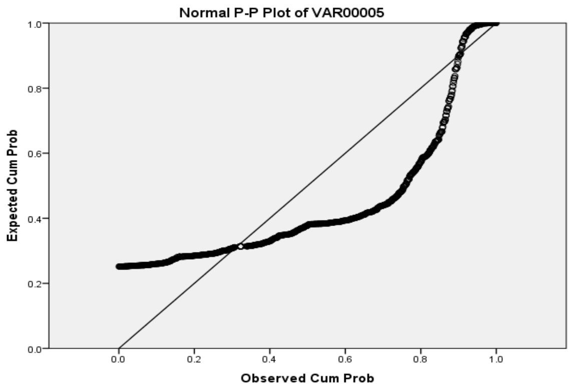

Fig. 8 shows the P-Plot of VAR00005, non-linear relationship exists, it is U-shaped type of relationship, data cannot be captured by just correlation and Regression

References

[1] A.F. Shapiro, "Fuzzy Regression Models", ARC 2005.

[2] O. Berman, F. Zahedi, & K.R. Pemble,, A decision model and support system for the optimal design of health information networks, IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews, 2001), 31(2): 146-158.

[3] J.M. Corrigan, & J.B. Martin,, Identification of Factors Associated with Hospital Readmission and Development of a Predictive Model, Health Services Research (1992), 27(1): 81-101.

[4] E. Demir, T. Chaussalet, H. Xie, & P.H. Millard, Modelling risk of readmission with phase- type distribution and transition models, IMA Journal of Management Mathematics (2008), 20(4): 357-367.
[5] M. Hensher, N. Edwards, & R. Strokes, International trends in the provision and utilization of hospital care, Bio Medical Journal (1999), 319(7213): 845-848.
[6] King's Fund Patients at risk of re-hospitalisation (PARR) case finding tool. Available at http://www.kingsfund.org.uk current projects/predictive risk/index.html, (2006).
[7] E.R. Marcantonio, S. McKean, M. Goldfinger, S. Kleefield, M. Yurkofsky, & T.A. Brennan, Factors Associated with Unplanned Hospital Readmission among Patients 65 Years of age and Older in a Medicare Managed Care Plan, The American Journal of Medicine (1999), 107(1): 13-17.
[8] E. Demir, T. Chaussalet, H. , Xie, & P.H. Millard, Modelling risk of readmission with phase- type distribution and transition models, IMA Journal of Management Mathematics (2008), 20(4): 357-367.
[9] P. Nagar, & S. Srivastava, Adaptive Fuzzy Regression Model for the Prediction of Dichotomous Response Variables using Cancer Data: A Case Study, Journal of Applied Mathematics and Informatics (JAMSI, 2008), 4(2): 183-191.

.
.