A Self Modified Mathematical Model for Interpolation of Disease Possibility in a Person Having Some Risk Factors

By

Syed Mohd Haider Zaidi

ISSN 0970-4973 Print
ISSN 2319-3077 Online/Electronic

Index Copernicus International Value
IC Value of Journal 4.21 (Poland, Europe) (2012)
Global Impact factor of Journal: 0.587 (2012)
Scientific Journals Impact Factor: 2.597

J. Biol. Chem. Research
Volume 29 (2) 2012 Pages No. 357-360

Journal of Biological and Chemical Research

An International Peer reviewed Journal of Life Sciences and Chemistry

Indexed, Abstracted and Cited: Index Copernicus International (Europe), Polish Ministry of Science and Higher Education (Poland, Europe) Research Bible (Japan), Scientific Journals Impact Factor Master List, Directory of Research Journals Indexing (DRJI), Indian Science. In, Database Electronic Journals Library (Germany), Open J-Gate, J Gate e-Journal Portal, Info Base Index, International Impact Factor Services (IIFS) (Singapore), Scientific Indexing Services (USA), International Institute of Organized Research (I2OR), Eye Source and

citefactor.org journals indexing

Directory Indexing of International Research Journals

Published by Society for Advancement of Sciences®

J. Biol. Chem. Research. Vol. 29, No. 2: 357-360 (2012)
(An International Peer reviewed Journal of Life Sciences and Chemistry)
Ms 30/2/230/2012

ISSN 2319-3077 (Online/Electronic)
ISSN 0970-4973 (Print)

http:// www.sasjournals.com
http:// www.jbcr.in
jbiolchemres@gmail.com
info@jbcr.in

RESEARCH PAPER

Received: 07/11/2012 **Revised: 03/12/2012** **Accepted: 04/12/2012**

A Self Modified Mathematical Model for Interpolation of Disease Possibility in a Person Having Some Risk Factors

Syed Mohd Haider Zaidi
Department of Statistics, Shia PG College, Sitapur Road, Lucknow – 226020, U.P., India

ABSTRACT

The techniques mostly used for interpolating disease possibility are logistic regression analysis and proportional hazards regression which generates the risk function with the help of some risk factors. But these techniques are not self modifiable and a complete overhauling is required after finding a new risk factor. The paper is introducing a self modifiable mathematical model for interpolation of disease possibility with the help of measurement theory.
Keywords: Mathematical Mode, Interpolation and Disease.

INTRODUCTION

The technique mostly used for interpolating disease possibility is logistic regression analysis which generates the risk function with the help of some risk factors. Many studies of this type are available in literature. For example in the study of Pullinger et. al (1993), logistic regression analysis was used to compute the odds ratios for 11 common occlusal features for asymptomatic controls with five temporomandibular disorder groups.

Another commonly used model to estimate disease risk is a proportional hazards regression, in which we study the impact of a risk factor on the occurrence of the disease with the time. For example, Arnlov et. Al. (2010) considers the effect of body mass index (BMI) and metabolic syndrome on the development of cardiovascular disease and death in middle-aged men.

But all the above techniques are not self modifiable and a complete overhauling is required after finding a new risk factor. Some more trials in this direction had been done using advanced methods of mathematical modeling. On such is provided in the paper of

Soonthornphisaj N. et. al. (2009), where a self-tuning of membership functions for fuzzy logic is proposed for medical diagnosis. The provided algorithm uses decision tree as a tool to generate three kinds of membership functions which are triangular, bell shape and Gaussian curve. The system can automatically select the best form of membership function for the classification process that can provide the best classification result. The system can create various membership functions using learning algorithm that learns from the training set.

A SELF MODIFIED FUNCTION

The method introduced in our paper used the results of Chen and Otto [1995] to construct membership functions of fuzzy sets using interpolation and measurement theory. Measurement theory provides a mathematically axiomatic method to construct membership function for a variable. The same theory can be applied here.

Let we define the disease risk function as $m\left(x_i\right)$,

Where x_i $\left(i=1,2,...n\right)$ denotes the intensity of i^{th} risk factor. And its value lies between 1 and 0 accordingly. *e.g* for lung cancer, smoking is a risk factor and its value varies from 1 to 0 for heavy to mild smoking.

The disease risk function to measure disease possibility is usually monotonic and convex and grows as number of risk factors increase. We define it in the range [0, 1] using the function-convex property :

$$m\left(\lambda x_i+\left(1-\lambda\right)x_j\right)\geq \min\left\{m\left(x_i\right),m\left(x_j\right)\right\} \qquad \text{......(1)}$$

where $\lambda\in\left[0,1\right]$ and $x_i,x_j\in R$.

Now we have desired properties of interpolated risk function in the form of axioms

Axiom 1 : A risk function is a numerical scale bounded in [0,1].

Axiom 2 : The rate that a modeller changes risk for values in a domain is continuous.

These axioms form a precise statement of what we seek from an interpolation function. We seek to find

$$m:R\rightarrow\left[0,1\right] \qquad \text{......(2)}$$

such that

1. $m\left(x\right)\in\left[0,1\right]\quad \forall x,$
2. *m* is differentiable
3. $m\left(x_i\right)=m_i$ for a finite set of known pairs $\left\{\left(x_1,m_1\right),...,\left(x_n,m_n\right)\right\}$
4. If the set of known pairs $\left\{\left(x_1,m_1\right),...,\left(x_n,m_n\right)\right\}$ is a convex set, then *m* is convex.

This problem statement is related to a calculus of variation problem, which is a method to determine a function that satisfies some constraints and simultaneously minimizes an "energy" function.

Now using calculus of variation approach we can develop an algorithm for finding a risk polynomial which is a second degree Bernstein polynomial.

The complete proposed algorithm to fit a risk function to a set of risk data $\{(x_1, m_1), \ldots, (x_n, m_n)\}$ is as follows :

1. First determine the slope t_i at each known point (x_i, m_i). To do this, first define

$$s_i = \frac{(m_i - m_{i-1})}{(x_i - x_{i-1})} \qquad \text{......(3)}$$

The following properties of t_i must be met :

(a) t_i must be consistent with the monotonicity and convexity of the piecewise linear function determined by the data points $\{(x_{i-1}, m_{i-1}), (x_i, m_i), (x_{i+1}, m_{i+1})\}$

(b) t_i must vary continuously with respect to changes in s_i and s_{i+1} when the signs of s_i and s_{i+1} agree.

(c) Only one knot should be required between two data points. This is to minimize the complexity of the algorithm.

The construction of t_i given below satisfies the above properties.

* If $s_i s_{i+1} \le 0$, we set $t_i = 0$. This guarantees that local extrema of the data points are assigned slopes 0. This also segments the entire data into monotonic subsets.

* If $|s_i| > |s_{i+1}| > 0$, we extend a line through $\vec{d}_i = (x_i, m_i)$ with slopes s_i until it intersects the horizontal line through $\vec{d}_{i+1} = (x_{i+1}, m_{i+1})$ at the point $\vec{b} = (b_x, m_{i+1})$.We then define

$$c_x = \frac{b_x + x_{i+1}}{2} \qquad \text{......(4)}$$

Which is the abscissa of point $\vec{c}$. Slope t_i at (x_i, m_i) is defined as

$$t_i = \frac{m_{i+1} - m_i}{c_x - x_i} \qquad \text{......(5)}$$

Note that

$$c_x > \frac{x_i + x_{i+1}}{2} \qquad \text{......(6)}$$

* If on the other hand, $0 < |s_i| < |s_{i+1}|$, we reverse the above procedure by extending the line through (x_i, m_i) with slope t_{i+1} until it intersects the horizontal line through (x_{i-1}, m_{i-1}) at the point $\vec{b} = (b_x, m_{i-1})$. Then we set

$$c_x = \frac{b_x + x_{i-1}}{2} \quad \text{and} \quad t_i = \frac{m_i - m_{i-1}}{x_i - c_x} \qquad \text{......(7)}$$

The end point slopes t_0 and t_n are set to zero explicitly.

2. We now insert a knot point between each x_i and x_{i+1} and fit a Bernstein polynomial to the $2n$–1 data points.

Proceeding in this way we can evaluate the risk function.

ACKNOWLEDGEMENTS
Author is grateful to Principal Shia P.G. College, Lucknow, U.P., India for providing all facilities during course of present investigation.

REFERENCES
A.G. Pullinger, D.A. Seligman, J.A. Gornbein (1993) "A Multiple Logistic Regression Analysis of the Risk and Relative Odds of Temporomandibular Disorders as a Function of Common Occlusal Features" Journal of Dental Research, Volume: 72 issue: 6, pp. 968-979

Johan Ärnlöv, Erik Ingelsson, Johan Sundström and Lars Lind (2010), "Impact of Body Mass Index and the Metabolic Syndrome on the Risk of Cardiovascular Disease and Death in Middle-Aged Men." Juornal of American Heart Association, Volume 121, Issue 2 pp. 121; 230-236,

Nuanwan Soonthornphisaj, Pattarawadee Teawtechadecha (2009) "A Self-tuning of Membership Functions for Medical Diagnosis" International Conference on Enterprise Information Systems (ICEIS) Enterprise Information Systems pp. 275-286.

Chain JE, Otto KN (1995) "Constructing membership functions using interpolation and measurement theory", Fuzzy Sets and Systems 73(3) pp.313-327a1. "Chapter - Methods of Clinical Epidemiology" Book- Springer Series on Epidemiology and Public Health, pp. 39-49

Corresponding author: Dr. Syed Mohd Haider Zaidi, Department of Statistics, Shia P.G. College, Sitapur Road, Lucknow – 226020, U.P., India
Email: syedmohdhaiderzaidi@yahoo.com