Saudi Journal of Biological Sciences (2016) 23, S11–S15

King Saud University

Saudi Journal of Biological Sciences

www.ksu.edu.sa
www.sciencedirect.com

ORIGINAL ARTICLE

A new mathematical evaluation of smoking problem based of algebraic statistical method

Maysaa J. Mohammed [a], Isamiddin S. Rakhimov [a], Mahendran Shitan [b], Rabha W. Ibrahim [c,*], Nadia F. Mohammed [a]

[a] Department of Mathematics, Faculty of Science, Universiti Putra Malaysia, Malaysia
[b] Institute for Mathematical Research (INSPEM), Malaysia
[c] Faculty of Computer Science and Information Technology, University Malaya, Malaysia

Received 22 April 2015; revised 19 July 2015; accepted 25 August 2015
Available online 1 September 2015

KEYWORDS

Markov basis;
Contingency tables;
Entropy

Abstract Smoking problem is considered as one of the hot topics for many years. In spite of overpowering facts about the dangers, smoking is still a bad habit widely spread and socially accepted. Many people start smoking during their gymnasium period. The discovery of the dangers of smoking gave a warning sign of danger for individuals. There are different statistical methods used to analyze the dangers of smoking. In this study, we apply an algebraic statistical method to analyze and classify real data using Markov basis for the independent model on the contingency table. Results show that the Markov basis based classification is able to distinguish different date elements. Moreover, we check our proposed method via information theory by utilizing the Shannon formula to illustrate which one of these alternative tables is the best in term of independent.

1. Introduction

US Department of Health and Human Services (2014) showed that smoking is responsible for roughly 440,000 deaths each year in the USA. Studies about the effects of smoking were done only on men in 1964 but later on women were included as well when some cigarette companies stated that smoking keeps girls and women thin. It can be noticed that health issues caused by smoking, no longer take into consideration, although most of people are familiar with the smoking consequences. Women who smoke significantly increase the danger of developing heart disease (the leading killer among women) and stroke. The danger increases with the number of cigarettes smoked and the length of time a woman has been smoking, but even people who smoke less than 5 cigarettes a day can have heart and blood vessel diseases. Although most of the women who die of heart disease are past menopause, smoking increases the danger additionally in younger women than in older women. Studies propose that smoking cigarettes increase the danger of heart disease even more among younger women who also take attractive birth control pills. Women who smoke, exclusively after going through menopause, have lower

* Corresponding author.
E-mail address: rahaibrahim@yahoo.com (R.W. Ibrahim).
Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.sjbs.2015.08.015
1319-562X

bone density (thinner bones). This means they have a higher danger of broken bones, counting hip fracture, than women who do not smoke. They could likewise be at higher danger of receiving rheumatoid arthritis and cataracts (clouding of the lenses of the eyes), as well as age-related macular degeneration, which can cause blindness. Tobacco habit can harm a woman's reproductive health. Women who smoke are more likely to have trouble getting pregnant. Smokers incline to be younger at the start of menopause than non-smokers and may have more unpleasant symptoms while going through menopause.

Nearly most of the people know how harmful it is being addictive smokers, however, few of them really recognize its risks. Some smokers convince themselves that they are social smokers or smoke only outdoors to be distinguished as non-smokers (Harris et al., 2002), and they can control smoking since smoking is not a habit for them. Stopping smoking is not that easy, yet being addictive does not mean a smoker does not have the ability to quit. The process of giving up smoking gets complicated, especially when people get older and older and the condition becomes sometimes irreversible. Some succeed from the early try while others give out more than one try; notwithstanding, some fail to quit smoking (Rollins et al., 2002). Moran et al. (2004) perceived that smokers start having a cigarette in occasional cases which means not daily and by the time they grow up, it gets irregular; then being involved in being motiveless to giving up; finally they reach a higher level (Moran et al., 2004). Due to several negative effects associated with this problem, many of the studies and research have emerged to study this problem in many respects, as in the above studies. This led to the attention of many researchers of the problem and the surrounding circumstances.

In this paper, we impose the algebraic statistical method using Markov basis for the independent model to find alternative data models that simulate the original data models and maintain the same statistical and mathematical qualities of the original data. On the other hand, our method is to find alternative data tables depending on the original data by using Markov basis for independent model. Moreover, we choose the best table in term of independent that liaises the qualities and characteristics is the same of the original data and the value of information on this table is more than the original data by using (information theory) as a measure for the independent. Due to the importance of passive smoking and its impact on all groups of society, we have selected data related to clarify the relationship between sex and smoking status (see Table 1).

Table 1 Contingency table for relationship between sex and smoking status.

Smoking status	Sex		
	Man	Woman	Total
Non smoker	10	6	16
Light smoker	8	8	16
Heavy smoker	4	8	12
Total	22	22	44

2. Material and methods

In 2007, Hannelore (2007) used Breathing test model. The data disquiet the relations among smoking position and breathing test outcomes for employers (under 40 years old) in a certain industrial plant. His study can be extended by presenting a further variable, namely the age. Observed are not the only employers under 40 years, but likewise employers from the age group 40–59. There are no restraints on the rows and column totals, and a simple model is that the count in the (i, j) cell y_{ij} is an understanding of a Poisson variable with expectation μ. The consequential likelihood is

$$L(\mu) = \prod_{i,j} \frac{\mu_{ij}^{y_{ij}}}{y_{ij}!} \exp(-\mu_{ij}) \tag{1}$$

where μ is the vector of the expectations.

In 2013, Jolanta et al. (2013) imposed the relationship between Peripheral vascular and smoking, and that smoking contributes significantly to the increase of cardiovascular disease and may increase the risk of holding such kinds of diseases by using the prevalence ratio calculation for cross section study.

$$\text{Prevalence} = \frac{\text{All new and preexisting cases during a given time period}}{\text{population during the same time period}} \times 10^n$$

The odds ratio (OR) is calculated as:

$$OR = \frac{a/b}{c/d} = \frac{ad}{bc} \tag{2}$$

In 2014, Suzan et al. (2014) found the relationship between the effect of smoking and disease, incontinence and the impact of smoking with the disease using hypothesis testing depending on the Fisher exact test. The study has shown that smoking has to do with the impact of a negative increase in case of this disease. The P-value calculated by formula of Fisher's exact test as:

$$P\frac{\left(\binom{a+b}{a}\binom{c+d}{c}\right)}{\binom{n}{a+c}} = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{a!b!c!d!n!} \tag{3}$$

In 2015, Anthony et al. (1999) studied that smoking during pregnancy has harmful effects and a large impact on the mother's weight at birth and also has significant side effects in children which may cause a respiratory problem and this leads to the lungs not working well at birth. Thus, in their studies, they have selected a sample of mothers either directly during a hospital admission or within 24 h of birth. The sample was taken from 100 mothers of non-smokers and 100 mothers who smoked 1–10 cigarettes per day and 100 mothers who smoked an average of more than 10 cigarettes per day. The aim of the study is to investigate the effects of smoking on lung function in newborns by knowing whether smoking during pregnancy affects or has to do with the functions of the lung in newborns. Using test hypotheses in statistics depending on chi square test. The formula of chi square test is:

$$X^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}, \tag{4}$$