Statistical Study On Some Eye Diseases (Such As Trachoma And Retinitis Pigmentosa)

Zahir Abdul Haddi Hassan

Mathematical department – College of education / Ibn Hayyan – Babylon university

Abstract

This paper deals with analysis of the effect of age on the incidence of some eye diseases such as patients with trachoma and inflammatory of the retina pigmentosa to stand on some important reasons that lead to the possibility of injury by avoiding the injury at an early stage before they escalate . The researcher to adjust the agenda approved a two-way for the classification of the data under the assumption of certain of the purpose of analysis and then access to the best model representing such data .

الخلاصة

يتناول هذا البحث تحليل تأثير العمر على الإصابة ببعض أمراض العيون مثل مرضي التراخوما والتهاب الشبكة الصباغي لمعالجتها قبل استفحالهما ، وقد قام الباحث بتوفيق جدول توافق ذي اتجاهين لتصنيف البيانات تحت فرضية معينة لغرض تحليلها ومن ثم الوصول الى أفضل نموذج يمثل تلك البيانات .

Introduction and aim of the research:

The eye disease is the most common disease among the people and the impact and impinge on their lives as it affects one of the most important senses and the blessings that God bestowed (the Almighty) in humans, namely, the grace of view and despite the scientific progress in various fields of life, particularly health and specifically where he's aware of Medicine after the eye made significant strides and reached by alleviating the suffering of people with eye disease in all its forms can not be found for the eradication of these diseases dramatically due to numerous factors and influences some of which directly affect and influence each other through interaction with others. Therefore, this research aims to determine the effect of age on worker injury patients, as trachoma and inflammatory of the retina pigmentosa of the eye diseases common among the people .

Research Sample :

Sample was obtained from Ibn al-Haytham Hospital records for 2001, and the size N = 1321 patients, were classified according to:

1 - Type of disease: which is here as a dependent variable:
 a - Trachoma .
 b - Inflammatory of the retina pigmentosa .
2 - Age in years: It has been divided into three categories:
 a – Less than 15 .
 b - 15 - 44 .
 c - 45 and more .

Where the number of patients within the age groups above as follows:

Age group (less than 15) 426 with 32.248 %
Age group (15 - 44) 450 with 34.065 %
Age group (45 and more) 445 with 33.687 %

The data has been put in a table of two-way contingency table .

Table (1)
Two-way contingency table
For the data of patients of trachoma and inflammatory of the retina pigmentosa

Age / Disease	less than 15	15 - 44	45 and more
Trachoma	231	230	210
inflammatory of the retina pigmentosa	195	220	235

Log – Linear Model : [Altohafy, Omar Hashim Abdul-Hamid, **2000**]
The estimate of expected value for the cell lies in row i and column j where i = 1 , … , r , j = 1 , … , c , Expressed as the following :

$$mij = \frac{(xi.)(x.j)}{N} \text{...............} \quad (1)$$

Where : N is the sum of observed frequencies $N = x_{..}$
$x_{i.}$ is the sum of observed frequencies in row i .
$x_{.j}$ is the sum of observed frequencies in row j .
The log of expected value in cell (i , j) in contingency table I x J is written as :

$$\text{Log } mij = u + u_1 i + u_2 j + u_{12} ij \quad \text{……………} (2)$$

Where :

$$u = \frac{1}{rc} \sum_{i=1}^{r} \sum_{j=1}^{c} Log\, mij \quad \text{....................} (3)$$

and : u_{1i} is the effect of first variable which lies in row i .
u_{2j} is the effect of second variable which lies in column j .
u_{12ij} is the effect of first and second variable which expressed the correlation between them .
The model (2) is called saturated model in two-dimensional contingency table , because it contain all the variables and its interactions from all degrees , starting with the effect of first variable u_{1i} , the effect of second variable u_{2j} and the interaction between them u_{12ij} . We can write the saturated model for all multi-dimensional contingency tables .
The log- linear model is called hierarchical model if two conditions are achieved:
1. In the event that any term of the range of u is equal to zero, then all upper commensurate terms with this term must equal to zero too.
2. In the event that any term of the range of u is not equal to zero, then all the commensurate terms with it that least than it in rank should not be equal to zero too , which mean they must be appear in the model.
For the purpose of access to the best model represents the data above, we would be passing several steps , first we have to obtain the expected estimates of the observed

data , here we will use the direct estimation method, which mentioned in the equation (1) above, where the marginal totals $x_{i.}$ and $x_{j.}$ are :
$x_{1.}$ = 671 sum of frequencies in the first row .
$x_{2.}$ = 650 sum of frequencies in the second row,
$x_{.1}$ = 426 sum of frequencies in the first column,
$x_{.2}$ = 450 sum of frequencies in the second column ,
$x_{.3}$ = 445 sum of frequencies in the third column ,
N = 1321 sample size .
After calculating the expected frequencies , the results as in table (2) :

Table (2)
The expected value for the independent model (1 , 2)

Model : Log mij = $u + u_1 i + u_2 j$		
Expected values matrix		
216.39	228.6	226.06
209.61	221.423	218.96

Now , we calculate the value of Person statistic χ^2 and Likelihood – ratio statistic G^2 , then we will compare them with the table value of χ^2 to know if the above independent model represent the data or not , that is if the calculated value of χ^2 and G^2 less than the table value of χ^2 then the calculated value will be morale and the independent model represents the data , but if the calculated value greater than the table value of χ^2 then we have to try with another model .
We can calculate χ^2 and G^2 by using the two following formulas :

$$\chi^2 = \sum_{i,j} \frac{(xij - mij)^2}{mij} \quad \text{......................} (4)$$

$$G^2 = 2\sum_{i,j} xij \ Log \ \frac{xij}{mij} \quad \text{...................} (5)$$

The results are : $\chi^2 = 4.337$ and $G^2 = 0.923$.
The degree of freedom of above independent model is 2 , which it known from following table : [Kazaz, Qutaiba Nabil Nayef, **2001**]

Table (3)
Degrees of freedom of saturated and non-saturated log – liner model for two-dimensional contingency table

Terms of u	Degrees of freedom
u	1
u_1	r -1
u_2	c -1
$u_{1,}u_2$	(r – 1) (c – 1)
Sum	r c

So , by comparing the calculating values of χ^2 and G^2 with the table value of χ^2 with 2 degrees of freedom under 0.05 level of morale which equal to $\chi^2 = 5.99$, we get that the calculating value less than the table one , that's mean the calculating value is morale and the independent model well good represents the data .

1. The common mean term u [Christensen, R. , Everitt, B.S. ,**1997**]:

$$u = \frac{1}{rc}\sum_{i=1}^{r}\sum_{j=1}^{c} Log\, mij \ldots\ldots\ldots\ldots\ldots\ldots.. \quad (6)$$

$$= 2.3426$$

2. The main effects terms for two variables :

A – Type of disease : (trachoma and inflammatory of the retina pigmentosa) :

We can find it from the formula [J.E. Figueroa-Lopez, **2009**] :

$$u_{1i} = \frac{1}{c}\sum_{j=1}^{c} Log\, m_{ij} - u \ \ldots\ldots\ldots\ldots\ldots.. (7)$$

Table (4)

The main effect term of the first variable with its levels u_{1i}

Disease type	Trachoma	inflammatory of the retina pigmentosa
level	0.0069	- 0.0069

B – The age (according to above age groups) :

We can find it from the formula [J.E.Figueroa-Lopez,**2008**]:

$$u_{2j} = \frac{1}{r}\sum_{i=1}^{r} Log\, m_{ij} - u \ \ldots\ldots\ldots\ldots\ldots.. (8)$$

Table (5)

The main effect term of the second variable with its levels u_{2j}

¶ Age group	15-44	44-64	64 and more
Level	-0.0165	0.006	0.0105

Note from the tables above that the largest main effects is the disease of trachoma in the age group (64 +) and this means that this disease most affects older age groups as a result of stress experienced by the eye in young period and not prevent it by checks frequent and require a comprehensive awareness of this age group of the risks that may affect their sense of sight as a result of this neglect, and that this awareness may lead to prevention and to reduce significantly the development of this disease [Yoshimasa Tsuruoka, **2009**].

Recommendations and proposals:

Recommend to the researcher and propose to the authorities responsible to educate young people a comprehensive alert (before it is too late) to the risk of this disease and the need for early screening and avoid repeated infection.

References

Altohafy, Omar Hashim Abdul-Hamid ,2000, "Contribution to study the relationship between the the asymmetric analysis and the log- linear model ",

Master in Statistics, Faculty of Business and Economics, University of Baghdad. Christensen, R. ,1997, "Log-Linear Models and Logistic Regression",Springer-Verlag Inc. New York, New York, USA.

Everitt, B.S. ,1977, "The Analysis of Contingency Tables", John Wiley & Sons, Inc., New York.

Figueroa-Lopez, J.E., 2009 , " Nonparametric estimation for levy models based on discrete- sampling", IMS Lecture Notes-Monograph Series. Optimality: The Third Erich L. Lehmann Symposium, 57:117-146.

Figueroa-Lopez, J.E.,2008, " Small-time moment asymptotics for levy processes",Statistics and Probability Letters, 78:3355-3365, .

Kazaz, Qutaiba Nabil Nayef , 2001,"Analysis of classification data in the contingency tables by using the biaz method ", Master in Statistics, Faculty of Business and Economics, University of Baghdad.

Yoshimasa Tsuruoka , Jun'ichi Tsujii and Sophia Ananiadou tochastic, **2009**"Gradient Descent Training for 1-regularized Log-linear Models with Cumulative enalty Proceedings of the 47th Annual Meeting of the ACL and the 4th IJCNLP of the AFNLP" , pages 477–485,Suntec, Singapore.