



THE 84TH
INTERNATIONAL
ELECTRONIC
CONFERENCE

ABSTRACTS

Contributed Papers

September 2-8, 1984
Keio Plaza Inter-Continental Hotel
Tokyo, Japan

AN APPLICATION OF HIERARCHICAL PATH MODEL TO INTERNATIONAL CANCER MORTALITY AND FOOD CONSUMPTION DATA

H. Takagi*, Y. Inaba**, and H. Yanai***

* St. Luke's College of Nursing

** Juntendo University

***Chiba University

Path analysis is originally devised by Wright (1921), and developed by Blalock (1964) for nonexperimental sciences. Recently, an algorithm was proposed by Takagi (1980), in order to build automatically a hypothetical causal model from the correlation matrix among an independent variable, say "effect", and dependent variables, say "causes". Since the procedure of selecting a variable in order and structuring the path model is similar to one of hierarchical methods of cluster analysis, i.e. nearest neighbour minimum method (single linkage), therefore we call it "Hierarchical Path Model (HPM)" here. The fundamental concepts are "causal chain" of variables and "direct effect" of a variable to another. Assume that p variables, say Z_1, Z_2, \dots, Z_p , cause Y directly, and that all variables are standardized. After another variable W is included into the model existed, the structure equation of Y is given by

$$Y = \sum_{i=1}^p D_i Z_i + D_W W + eU$$

where D_i and D_W represent direct effects (path coefficients) of $Z_i (i=1, 2, \dots, p)$ and W respectively, and U and e are a disturbance influencing Y and its weight respectively. If variable W correlates highly with one of p causal variables, say X , D_W is estimated approximately by

$$D_W = \frac{R_{YW} - R_{YX}R_{XW}}{1 - R_{XW}^2}$$

where R_{XW} is a correlation coefficient between variables X and W , and so on. The central procedure of HPM consists of detecting a pair of variables with the highest correlation, one is included in the model but another is not, evaluating the direct effect of selecting variable to Y , and linking it to the model. The procedure of HPM will be described simply.

Using international cancer mortality and food consumption data, we build a hypothetical causal model, estimate and evaluate the strength of each effect of causal variable to another. Moreover, we compare the results with those of other methods of multivariate analysis, especially multiple regression analysis. Through the comparison of the several results, it is suggested that HPM might be a powerful tool to describe the relationship among variables, and that the concurrent use of multivariate analysis could be helpful to interpret the results of the study, especially in the field of epidemiology.

(This study was supported partly by the Kagaku-kenkyu-hi Shourei-kenkyu A, the grant of the Ministry of Education, Japan, 1982)

REFERENCES

- Blalock, H.M. Jr. (1964). Causal inferences in nonexperimental research, the University of North Carolina, Chapel Hill, North Carolina.
 Takagi, H. (1980). An algorithm of causal modeling in epidemiological studies, Behaviormetrika 8, 41-55.
 Wright, S. (1921). Correlation and causation, Journal of Agricultural Research 20, 557-585.

AN EPIDEMIOLOGICAL ANALYSIS OF CANCER MORTALITIES
IN JAPAN BY THE QUANTIFICATION METHOD FOR THREE WAY DATA

H. Yanai¹⁾ Y. Inaba²⁾ H. Takagi³⁾

1) Chiba University 2) Juntendou University

3) St. Luke's College of Nursing

The present paper was made for the purpose of analysing cancer mortalities from 1958 to 1978 in Japan in order to obtain some clues which may lead to elucidating the unknown etiology regarding various malignant neoplasms, by developing a method called the quantification method for three way data introduced first by Takeuchi, Yanai & Mukherjee(1982). The fundamental idea of the method lies in finding weights $(a_1, \dots, a_p), (b_1, \dots, b_q)$ and (c_1, \dots, c_r) maximizing

$$\sum_i \sum_j x_{ij.} a_i b_j + \sum_j \sum_k x_{.jk} b_j c_k + \sum_i \sum_k x_{i.k} a_i c_k$$

$$\text{subject to } \sum_i x_{i..} a_i^2 = \sum_j x_{.j.} b_j^2 = \sum_k x_{..k} c_k^2 = 1$$

where x_{ijk} is the (i,j,k) element of the three way data, and

$$x_{ij.} = \sum_k x_{ijk}, \quad x_{.jk} = \sum_i x_{ijk}, \quad x_{i.k} = \sum_j x_{ijk}$$

$$x_{i..} = \sum_j \sum_k x_{ijk}, \quad x_{.j.} = \sum_i \sum_k x_{ijk}, \quad x_{..k} = \sum_i \sum_j x_{ijk}$$

Theoretically, the method introduced here can be said to be a special case of the generalized canonical correlation analysis introduced by Horst(1961). Of special interests is the fact that we can extract more than two factors from each of the three dimensions i, j and k .

The method was applied to the following two kinds of three way data regarding the record on cancer mortalities ; i.e., (i) selected 18 sites of cancer, calendar year(from 1958 to 1978) and 46 prefectures of Japan for male and female respectively, (ii) selected sites of cancer, calendar year and 18 age grade by five years classification. Furthermore, principal component analysis was applied to some kinds of two way data.

The quantification method based on the three way data set (i) revealed that mortality rates of lung and prostate cancers in males have been increasing year by year particularly in the south Kyushu district(southern part of Japan), and those of pancreas and bile duct cancers have been on the increase as the calendar years proceed on. Furthermore, analysis based on the data set (ii) revealed that mortality rates of lung and prostate cancers have been on the increase followed by the increment of the age at death. Principal component analysis followed by varimax rotation yielded the six factors which are suspected to have different etiological agents regarding some specified environmental factors. Of special interest is the fact that lung and liver cancers showed marked changes of factor loadings in the first factor, and that intestine and rectum cancers showed marked changes in the loadings of the third and fourth factors.

REFERENCES

- Takeuchi K., Yanai H. & Mukherjee B.N.(1982) The Foundation of Multivariate Analysis, Wiley Eastern
Horst P.(1961) Relations among m sets of measures, Psychometrika, 26,129-149
Inaba Y, Yanai H., Takagi H. & Yamamoto S.(1981) A study of the geographical pattern of cancer mortality for selected sites by means of factor analysis, Social Science & Medicine, Vol 15, NO 1, 233-244