(1/2)

NSC92-2119-M009-002-

92　12　01　　93　07　31

93　5　13

Application of Latent Variable Models in Evaluating Diagnostic Tests and Its Software Development

Abstract

In many medical studies, the definitive outcome is inaccessible due to cost, time and difficulty of measurement. A valid surrogate endpoint is then measured in place of the biologically definitive or clinically most meaningful endpoint. For example, biomarkers are very often used as surrogates of observing new cases of cancer in testing treatments for cancer prevention, where event rates are low and a long time may be needed to obtain definitive outcomes. Functional disability is commonly quantified as self-reported responses to a series of questions about difficulty performing tasks of routine living, because no obvious single measure of disability exists.

Latent variable models are recognized as an effective statistical tool for analyzing surrogates. Such models treat the unobserved definitive outcome as the one to be analyzed for relations with risk factors, and the measured surrogates as quantities that imperfectly determine the object of interest. Analysis based on latent variable models is parsimonious, explicitly recognize errors in measurement, and can give well-summarized inferences on the theory underlying the choice of surrogates.

The author has recently developed a very flexible latent variable model for analyzing surrogates. I also addressed issues

1

about selecting the number of dimensions needed to characterize the surrogates' distribution and diagnosing various latent variable model assumptions. In this project, I propose to extend the latent variable model developed by the author to evaluate the validity and reliability of diagnostic tests. This application is important because the model enables us to evaluate tests without knowing the gold standard. I also plan to incorporate the latent variable methodologies developed by the author into an easy-to-use statistical software capable of being used effectively by all levels of participants of data analysis.

Keywords: latent variable, reliability, software, surrogate endpoint, validity, web page.

In many medical studies, the definitive outcome is inaccessible. A valid surrogate endpoint is then measured in place of the biologically definitive or clinically most meaningful endpoint.

Models that permit exploration of relationships between unobservable variables and their surrogates are referred to as latent variable models. My research is focused on analyzing data collected in situations where investigators use multiple discrete indicators as surrogates, for example, a set of questionnaires. These models are called the latent class analysis (LCA) model (Lazarsfeld and Henry 1968, Goodman 1974). The basic LCA postulates an underlying categorical latent variable, and this latent variable can explain the association seen among measured items. Recently, several authors extended the LCA model to incorporate covariate effects on estimating the underlying mechanism (Dayton and Macready 1988, Bandeen-Roche et al. 1997), or on estimating measured

indicator distributions within latent classes (Melton et al. 1994). I and co-author proposed an LCA model that used covariates on describing distributions of both the underlying latent class and the measured indicators themselves (Huang and Bandeen-Roche 2003A). Significantly, we developed a model framework that guarantees identifiability of the two types of covariate effects.

In this project, I propose to extend the latent class model developed by the author to evaluate the validity and reliability of diagnostic tests. This application is important because the model enables us to evaluate tests without knowing the gold standard. I also plan to incorporate all the latent variable methodologies developed by the author into an easy-to-use statistical software capable of being used effectively by all levels of participants of data analysis. These methodologies include the proposed LCA mode (Huang and Bandeen-Roche 2004A), selecting the number of dimensions needed to characterize the surrogates' distribution (Huang 2004B), diagnosing various latent variable model assumptions (Huang and Bandeen-Roche 2004C), and evaluating tests without the gold standard (will be done in this proposal).

(http://www.stat.nctu.edu.tw/faculty/ghuang/software.htm)

3

1. **Huang GH**, Bandeen-Roche K (2004). Building an identifiable latent variable model

with covariate effects on underlying and measured variables. Accepted for publication in *Psychometrika*

Abstract:
In recent years, latent class models have proven useful for analyzing relationships between measured multiple indicators and covariates of interest. Such models summarize shared features of the multiple indicators as an underlying categorical variable, and the indicators' substantive associations with predictors are built directly and indirectly in unique model parameters. In this paper, we provide a detailed study on the theory and application of building models that allow mediated relationships between primary predictors and latent class membership, but also allow direct effects of secondary covariates on the indicators themselves. Theory for model identification is developed. We detail an Expectation-Maximization algorithm for parameter estimation, standard error calculation and convergent properties. Comparison of the proposed model with models underlying existing latent class modeling software is provided. A detailed analysis of how visual impairments affect older persons' functioning requiring distance vision is used for illustration.

2. **Huang GH** (2004). Selecting the number of classes under latent class regression: a factor analytic analogue. Accepted for publication in *Psychometrika*.

Abstract:
Recently, the regression extension of latent class analysis (RLCA) model has received much attention in the field of medical research. The basic RLCA model summarizes shared features of measured multiple indicators as an underlying categorical variable and incorporates the covariate information in modeling both latent class membership and multiple indicators themselves. To reduce complexity and enhance interpretability, one usually fixes the number of classes in a given RLCA. Often, goodness of fit methods comparing various estimated models are used as a criterion to select the number of classes. In this paper, we propose a new method that is based on an analogous method used in factor analysis and does not require repeated fitting. Two ideas with application to many settings other than ours are synthesized in deriving the method: a connection between latent class models and factor analysis, and techniques of covariate marginalization and elimination. A Monte Carlo simulation study is presented to evaluate the behavior of the selection procedure and compare to alternative approaches. Data from a study of how measured visual impairments affect older persons' functioning are used for illustration.

3. **Huang GH**, Bandeen-Roche K (2004). Diagnosis of latent class regression models. Submitted for publication.

Abstract:
Biomedical and psychosocial researchers increasingly utilize latent class regression (LCR) models to analyze relationships between measured multiple categorical outcomes and covariates of interest. In LCR, the multiple outcomes are summarized and their associations with risk factors are determined in a single modeling step. These models are parsimonious and can incorporate theory underlying the multiple response choices. However, these advantages come at the price of strong modeling assumptions which may critically influence analytic findings. Careful evaluation of model appropriateness is necessary. In this paper, we first develop Pearson and deviance residuals to assess overall fit of the LCR

model, and then propose reasonable methods for diagnosing various LCR's assumptions. The basic idea of the latter diagnostic method is to first estimate the underlying latent class, and then directly assess model assumptions within and across classes. This approach is proven to be valid in an asymptotic sense. An analysis of how measured health impairments affect older persons' functioning is used for illustration.

[1] Albert PS, McShane LM, Shih JH (2001). Latent class modeling approaches for assessing diagnostic error without a gold standard: with applications to p53 immunohistochemical assays in bladder tumors. *Biometrics* 57:610-619.

[2] Bandeen-Roche K, **Huang GH**, Munoz B, Rubin GS (1999). On determining risk factor associations with questionnaire outcomes: a methods case study. *American Journal of Epidemiology* 150:1165-1178.

[3] Bandeen-Roche K, Miglioretti DL, Zeger SL, Rathouz PJ (1997). Latent variable regression for multiple discrete outcomes. *Journal of the American Statistical Association* 92:1375-1386.

[4] Cutting LE, **Huang GH**, Zeger S, Koth CW, Thompson RE, Denckla MB (2002). Growth curve analyses of neuropsychological profiles in children with neurofibromatosis type 1: specific cognitive tests remain "spared" and "impaired" over time. *Journal of the International Neuropsychological Society* 8:838-846.

[5] Dayton CM, Macready GB (1988). Concomitant-variable latent-class models. *Journal of the American Statistical Association* 83:173-178.

[6] Goodman LA (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika* 61:215-231.

[7] Gurewitsch ED, Diament P, Fong J, **Huang GH**, Popovtzer A, Weinstein D, Chervenak FA (2002). The labor curve of the grand multipara: does progress of labor continue to improve with additional childbearing? *American Journal of Obstetrics and Gynecology* 186:1331-1338.

[8] **Huang GH** (2004B). Selecting the number of classes under latent class regression: a factor analytic analogue. Accepted for publication in *Psychometrika*.

[9] **Huang GH**, Bandeen-Roche K (2004A). Building an identifiable latent variable model with covariate effects on underlying and measured variables. Accepted for publication in *Psychometrika*.

[10] **Huang GH**, Bandeen-Roche K (2004C). Diagnosis of latent class regression models. Submitted for publication.

[11] **Huang GH**, Bandeen-Roche K, Rubin GS (2002). Building marginal models for multiple ordinal measurements. *Journal of the Royal Statistical Society, Series C (Applied Statistics)* 51:37-57.

[12] **Huang GH**, Klein R, Klein BEK, Tomany SC (2003). Birth cohort effect on prevalence of age-related maculopathy in the Beaver Dam Eye Study. *American Journal of Epidemiology* 157:721-729.

[13] **Huang GH**, Palta M, Allen C, LeCaire T, D'Alessio D for the Wisconsin Diabetes Registry (2004). Self-rated health in young persons with Type 1 diabetes in relation to risk factors from a longitudinal study. *American Journal of Epidemiology* 159:364-372.

[14] Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffer MW (1963). Studies of illness in the age. The index of ADL: a standardized measure of biological and psychosocial function. *Journal of the American Medical Association* 185:914-918.

[15] Klein R, Klein BEK, Tomany SC, Meuer SM, **Huang GH** (2002). Ten-year incidence and progression of age-related maculopathy: The Beaver Dam Eye Study. *Ophthalmology* 109:1767-1779.

[16] Lazarsfeld PF, Henry NW (1968). *Latent Structure Analysis*. New York: Houghton-Mifflin.

[17] Melton B, Liang KY, Pulver AE (1994). Extended latent class approach to the study of familial/sporadic forms of a disease: its application to the study of the heterogeneity of schizophrenia. *Genetic Epidemiology* 11:311-327.

[18] Piantadosi S (1997). Clinical Trials: A Methodologic Perspective. New York: Wiley.

[19] Qu Y, Tan M, Kutner MH (1996). Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. *Biometrics* 52:797-810.

[20] Rubin GS, Bandeen-Roche K, **Huang GH**, Munoz B, Schein OD, Fried LP, West SK (2001). The association of multiple visual

impairments with self-reported visual disability: SEE project. *Investigative Ophthalmology and Visual Science* 42:64-72.

[21] Statistical Science Inc. (1994). *S-PLUS Reference Manual*, Version 3.1, Seattle: Statistical Science; a division of Mathsoft, Inc.

[22] Uebersax JS, Grove WM (1993). A latent trait finite mixture model for the analysis of rating agreement. *Biometrics* 49:823-835