

From disability to not self-sufficiency: an application of latent variables models¹

*Dalla disabilità alla non-autosufficienza:
un'applicazione dei modelli a variabili latenti.*

Giorgio E. Montanari — Paolo Eusebi — M. Giovanna Ranalli
Dipartimento di Economia, Finanza e Statistica, Università degli Studi di Perugia
giorgio@stat.unipg.it - peusebi@stat.unipg.it - giovanna@stat.unipg.it

Riassunto: In questo lavoro si propone un'applicazione dei modelli a classi latenti per la quantificazione della non autosufficienza negli anziani, con particolare riferimento ad un'area territoriale dell'Italia centrale. A tal fine sono stati utilizzati i dati dell'indagine ISTAT Multiscopo sulle *Condizioni di salute e ricorso ai servizi sanitari 1999-2000*, relativi alle variabili che rilevano le difficoltà nelle attività quotidiane. La selezione del modello ha portato all'individuazione di quattro classi latenti che dividono gli anziani in *autosufficienti, con disabilità lieve, con non autosufficienza iniziale e non autosufficienti*. Successivamente, mediante un modello di Rasch per variabili ordinali è stato attribuito a ciascun anziano un punteggio sulla scala del tratto latente della non autosufficienza allo scopo di verificare la bontà della classificazione ottenuta.

Keywords: Categorical Variables, Latent Class Model, Rasch Model, Partial Credit Model, Item, Activities of Daily Living.

1. Introduction and data

In this work we study and quantify the phenomenon of not self-sufficiency in the elderly with a particular focus on the Italian administrative district of Umbria and neighboring provinces. Data come from the national survey on *Health Conditions and Appeal to Medicare 1999-2000*, conducted by the Italian National Institute of Statistics (ISTAT, 2002). The questionnaire was constructed accounting for the International Classification of Impairments, Disabilities and Handicaps (ICIDH) developed by the World Health Organization (WHO, 1980). The latter provides a conceptual framework for disability, which is described in three dimensions: Impairment, Disability and Handicap. In the context of health experience, an impairment is any loss or abnormality of psychological, physiological or anatomical structure or function. On the other side, a disability is any restriction or lack (resulting from an impairment) of ability to perform an activity in the manner or within the range considered normal for a human being.

In the national survey quoted above, a person is considered disabled if, excluding temporary limitations, he/she expresses the largest degree of difficulty in at least one of a set of functions concerning the ability of a person to accomplish everyday tasks such as getting washed and dressed, eating, walking, even with the aid of tools such as glasses, walking sticks, prostheses. Such functions are evaluated by mean of 16 questions (items) in the questionnaire that include the *Activities of Daily Living* (ADL; Katz et al., 1963). The items are reported in Table 1; they are all ordinal with categories increasing

¹ The present research is financially supported by the Region of Umbria.

with the difficulty of fulfilling the task and they are grouped according to four types of disabilities.

Clearly, even though an elder express the maximum degree of difficulty in fulfilling one of these tasks, and, therefore, he/she is labeled as “disabled”, it does not imply being not self-sufficient: a person may be unable to hear a TV show, but he/she can be perfectly self-sufficient. A person is defined as not self-sufficient when affected by a permanent disability or a chronic disease, that deprives his/her autonomy so much that he/she needs continuous personal assistance to complete basic everyday tasks. Not self-sufficiency is very important to assess for its consequences on Health system organization, policy making and funding.

To estimate the quantity and, therefore, the impact of not self-sufficient elderly using the ISTAT data, we employ a Latent Class Model (LCM; Lazarsfeld and Henry, 1968; Goodman, 1974): it finds homogeneous groups (latent classes) from multivariate categorical data. Homogeneity of the classes is defined in terms of a latent trait, which would hopefully be, in our case, connected to self-sufficiency. The application of this method, together with that of the Partial Credit Model (PCM; Masters, 1982) that gives a score to each elder according to that latent trait is reported in the following section.

Table 1: *Type of disability and corresponding Items surveyed.*

1. confinement	CHOM = confined to one's home	CSIT = confined to a wheelchair	CBED = confined to one's bed
2. difficulties in movements	DIST = longest walkable distance	STAIR = going up and down the stairs	STOOP = stooping down
3. difficulties in everyday activities and tasks	BED = getting in and out of bed	CHAIR = sitting and standing	DRESS = getting dressed and undressed
	BATH = taking a bath or a shower	WASH = washing one's face and hands	EAT = eating cutting one's food
	CHEW = chewing		
4. sensory deprivation	TV = hearing a TV show	SIGHT = seeing and recognizing a friend	SPEE = speaking

2. Latent variables models

Let us consider a LCM with a single latent trait θ and J ordinal items Y_j with $j = 1, 2, \dots, J$. Latent classes are denoted by $c = 1, 2, \dots, C$. The basic assumption of the LCM is that items are independent of one another within latent classes (Goodman, 1974). A LCM with a single latent variable can be defined as

$$P(h_1, \dots, h_j, \dots, h_J) = \sum_{c=1}^C \Pr(\theta = c) \prod_{j=1}^J \Pr(Y_j = h_j | \theta = c) \quad (1)$$

where $h_1, h_2, \dots, h_J$ are any set of responses to the J items.

In a LCM the latent trait is described by a certain number of classes with unknown sizes. Successive LCMs are fitted to the data with increasing number of classes until the simplest model is found that provides an adequate fit. In a traditional LCM analysis cases are classified into the appropriate latent classe as follows. Let us consider two ordinal items; for any given response pattern ($Y_l = h_l, Y_m = h_m$), estimates for the posterior

membership probabilities are obtained using Bayes theorem as follows:

$$\hat{P}_{clm}^{C|LM} = \hat{P}(c) \hat{P}(Y_l = h_l | c) \hat{P}(Y_m = h_m | c) / \sum_{c=1}^C \hat{P}(c) \hat{P}(Y_l = h_l | c) \hat{P}(Y_m = h_m | c), \quad (2)$$

where estimates of the parameters coming from equation (1) are inserted in the right-hand side. Cases are then assigned to the class for which the posterior membership probability takes the largest value.

We used a data set containing 3150 cases from Umbria and the neighboring provinces and 14 items (CHOM, CSIT and CBED are combined in one dichotomus item CONF). We fitted LCMs with WinMira software (Von Davier, 1995) and tested different LCMs with one up to six classes. We found an adequate classification with four classes (see Figure 1a that reports CAIC and BIC values for the six models fitted): as a result, about 61% of the observations belongs to the first class, 21% to the second one, 12% to the third one and the remaining 6% to the fourth one. People belonging to each class have been labeled in the following way (from the first one to the fourth one): self-sufficient, with light disability, with initial not self-sufficiency, not self-sufficient. Such interpretation of the latent classes comes from the estimated conditional probabilities of the degrees of difficulty of each item. To sketch the situation, Table 2 reports the conditional probabilities of the largest degree of difficulty for each item and latent class.

Table 2. *Conditional probabilities of the greatest difficulty category of each item*

	self-sufficient	with light disability	with initial not self-sufficiency	not self-sufficient
CONF	0.004	0.017	0.216	0.917
DIST	0.000	0.015	0.280	0.950
STAIR	0.000	0.007	0.192	0.917
STOOP	0.000	0.015	0.180	0.950
BED	0.000	0.003	0.075	0.942
CHAIR	0.000	0.000	0.014	0.757
DRESS	0.000	0.000	0.118	0.896
BATH	0.002	0.071	0.429	0.967
WASH	0.000	0.000	0.034	0.551
EAT	0.000	0.000	0.050	0.462
CHEW	0.001	0.003	0.014	0.094
TV	0.011	0.030	0.074	0.218
SIGHT	0.006	0.007	0.039	0.164
SPEE	0.001	0.003	0.010	0.080

Self-sufficient people show all almost zero values and, therefore, no limitation in all functions. In the second class, we find large probabilities (not reported here) of observing people with some difficulties in walking or taking a bath or a shower. People with an initial state of not self-sufficiency show some difficulty in almost all activities but are not usually confined. In the class of not self-sufficiency we have the largest probability of observing people who express the maximum level of difficulty for all activities, but chewing and those concerning sensory deprivation. It should be noted that these four items, indeed, showed the least discrimination power among all items and, if removed

from the analysis, do not affect the classification.

Scales summarizing information from item responses are also often used to infer the amount of the latent trait. When polytomously scored items are considered, a PCM can be used within the broader family of Rasch Models (Rasch, 1960) for ordinal items. The response to item Y_j depends on the amount of the latent trait θ and the difficulty parameter β_{jh_j} of the specific category h_j of item Y_j : $P(Y_j = h_j) = \exp(\theta - \beta_{jh_j}) / \sum_l \exp(\theta - \beta_{jl_j})$. The values taken by the Rasch estimates are indeed consistent with the latent classes membership, in that largest values of the Rasch person parameter correspond to observations in the fourth class and small values to observations in the first class. Figure 1b reports the percent conditional distributions of the Rasch scores for each class. They well support the proposed classification of elderly with respect to self-sufficiency.

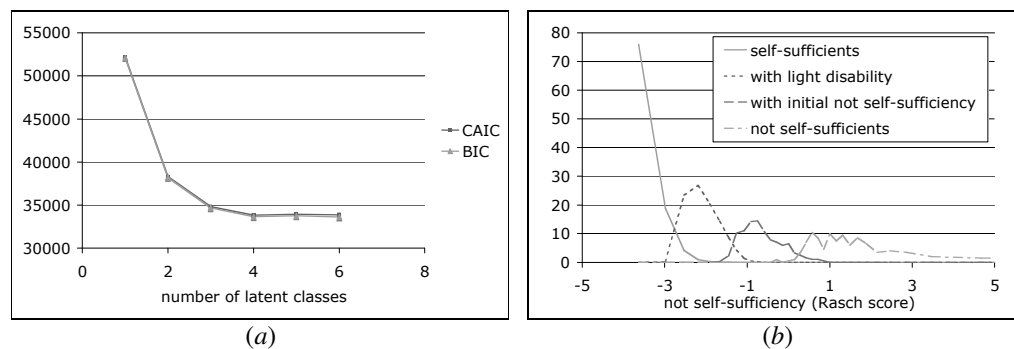


Figure 1. (a) Model selection through information criteria. (b) Distribution of Rasch scores within latent classes

References

- Goodman, L. A. (1974) Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, 61, 215-231.
- ISTAT (2002) Le condizioni di salute della popolazione – Indagine Multiscopo sulle famiglie “Condizioni di salute e ricorso ai servizi sanitari. Anni 1999-2000, *Informazioni* n.12, Roma.
- Lazarsfeld P.F., Henry N.W. (1968) *Latent structure analysis*. Boston, Houghton Mifflin.
- Katz S., Ford A. B., Moskowitz R. W., Jackson B. A., Jaffe M. W. (1963) Studies of illness in the aged. The Index of ADL: a standardized measure of biological and psychosocial function. *Journal of the American Medical Association*, 185, 914-919.
- Masters, G. (1982) A Rasch model for partial credit scoring, *Psychometrika*, 47, 149-175.
- Rasch, G. (1960) *Probabilistic Models for Some Intelligence and Attainment Tests*. The Danish Institute of Educational Research, Copenhagen.
- Von Davier M. (1995) *WinMira v.1.74 for Windows* Kiel: Institute for Science.
- World Health Organisation (WHO) (1980) *International Classification of Impairments, Disabilities and Handicaps*. World Health Organisation, Geneva.