An approach for the assessment of the Health-related Quality of Life based on multidimensional latent class Rasch models

Silvia Bacci*[1], Francesco Bartolucci*

*Department of Economics, Finance and Statistics - University of Perugia

Patient Reported Outcomes and Quality of Life

Paris, 4-5 July 2011

[1] silvia.bacci@stat.unipg.it

Outline

Health related Quality of Life (HrQoL)

HrQoL is

> *"the way, which according to health of a person influences his/her capacity to lead on physical and social normal activities."*

(Fayers and Machin, 2000)

- HrQoL is a **latent** construct
  - It is indirectly measured by a set of items;
  - **Item Response Theory (IRT) models** allow us to translate the qualitative information coming from the questionnaire in a quantitative measurement of the latent construct.
- HrQoL is a **multidimensional** construct
  - **Multidimensional extensions of IRT models** represent a suitable framework to take into account the correlation between the different dimensions of one or more latent traits.

Main aims of the study of HrQoL

1. Detecting **homogeneous classes of individuals** who have very similar latent characteristics (Goodman, 1974; Lazarsfeld and Henry, 1978).

   Detecting these classes of individuals can be, not only more realistic, but also more convenient for the decisional process because individuals in the same class will receive the same clinical treatment.

2. Studying the **correlation** between dimensions of HrQoL and between HrQoL and some psychopathological disturbs (e.g. anxiety, depression).

Aims of the contribution

1. Estimating a **multidimensional latent class (LC) Rasch model** (Bartolucci, 2007), so as to detect homogeneous classes of individuals;
2. Performing two tests of dimensionality and analysing the correlation between dimensions of HrQoL and between HrQoL and anxiety and depression.
   - Likelihood ratio (LR) test based on the multidimensional LC Rasch model that exploits the **discrete marginal maximum likelihood (MML)** approach;
   - Martin-Löf test based on the **conditional maximum likelihood (CML)** estimation method.
3. These aims are illustrated through an application to a real dataset concerning HrQoL in cancer patients.

The questionnaires

Data have been collected by using two questionnaires:

- HrQoL assessed by the "36-item short-form health survey" (SF-36) (Ware et al., 2002)
  - 9 latent dimensions: physical functioning, role functioning, bodily pain, general health, vitality, social functioning, role-emotional, mental health, health change;
  - 36 polytomous items;
  - dichotomization of response categories: $1 =$ presence of a symptom or limitation (related to a low level of HrQoL); $0 =$ absence of a symptom or a limitation (related to a high level of HrQoL).
- Anxiety and depression assessed by the "Hospital Anxiety and Depression Scale" (HADS) (Zigmond and Snaith, 1983)
  - 2 latent dimensions: anxiety and depression;
  - 14 polytomous items;
  - dichotomization of response categories: $1 =$ presence of anxiety (or depression); $0 =$ absence of anxiety (or depression).

Description of the dataset

- 275 oncological patients recruited from three different Italian centres
- response rate equal to 74%

	Entire sample	Respondents
Age (years)		
Mean	54.6	54.3
St. Dev.	13.4	11.5
Gender (%)		
Female	66.9	68.9
Male	33.1	31.1
Cancer diagnosis (%)		
Colon-rectum	24.4	23.9
Mammary	45.6	46.7
Uterine	4.1	3.8
Pulmonary	8.8	8.7
Prostate	4.1	3.8
Other	13.0	13.0
Size	275	203

Basic notation

- n: sample size ($n = 203$ in the application)
- k: number of binary items ($k = 50$)
- D: number of latent traits or dimensions ($D = 11$)
- $\mathcal{I}_d$, $d = 1, \dots, D$: subset of $\mathcal{I} = \{1, \dots, k\}$ containing the indices of the items measuring the latent trait of type d
- k_d: cardinality of $\mathcal{I}_d$, $k = \sum_{d=1}^{D} k_d$
- $X_i = 0, 1$: random variable corresponding to the response to item i
- β_i: difficulty of item i
- $\boldsymbol{\Theta} = (\Theta_1, \dots, \Theta_D)'$: vector of latent variables corresponding to the different traits measured by the test items
- $\boldsymbol{\theta} = (\theta_1, \dots, \theta_D)'$: one of the possible realizations of $\boldsymbol{\Theta}$
- δ_{id}: dummy variable equal to 1 if item i belongs to $\mathcal{I}_d$ and to 0 otherwise

Assumptions (1)

- Rasch model (Rasch, 1961)

$$\text{logit}[p(X_i = 1 \mid \Theta = \theta)] = \theta - \beta_i, \quad i = 1, \ldots, k$$

- Multidimensional Rasch model

$$\text{logit}[p(X_i = 1 \mid \boldsymbol{\Theta} = \boldsymbol{\theta})] = \sum_{d=1}^{D} \delta_{id}\theta_d - \beta_i, \quad i = 1, \ldots, k$$

  - Between-item multidimensionality (Adams et al., 1997): each item measures only one latent trait
  - Latent trait $\boldsymbol{\Theta}$ may be assumed as random rather than fixed and, in the random-effects approach, discrete rather than continuous

Assumptions (2)

- Multidimensional LC Rasch model
  - The random vector $\boldsymbol{\Theta}$ is assumed to have a discrete distribution with support points $\{\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_C\}$, i.e. the population is assumed to be composed by C subpopulations
  - The number C of latent classes is the same for each dimension
  - Manifest distribution of the full response vector $\mathbf{X} = (X_1, \ldots, X_k)'$:

$$p(\mathbf{x}) = p(\mathbf{X} = \mathbf{x}) = \sum_{c=1}^{C} p(\mathbf{x}|\boldsymbol{\zeta}_c)\pi_c$$

where $\pi_c = p(\boldsymbol{\Theta} = \boldsymbol{\zeta}_c)$ and (assumption of *local independence*)

$$p(\mathbf{x}|\boldsymbol{\zeta}_c) = p(\boldsymbol{X} = \mathbf{x}|\boldsymbol{\Theta} = \boldsymbol{\zeta}_c) = \prod_{d=1}^{D}\prod_{h=1}^{k_d} p(X_{dh} = x_{dh} \mid \Theta_d = \zeta_{cd}), \quad c = 1, \ldots, C$$

Maximum log-likelihood estimation

Let $n(\mathbf{x})$ denotes the frequency of the response configuration $\mathbf{x}$ in the sample and let $\boldsymbol{\eta}$ the vector containing all the free parameters. The log-likelihood may be expressed as

$$\ell(\boldsymbol{\eta}) = \sum_{\mathbf{x}} n(\mathbf{x}) \log[p(\mathbf{x})]$$

- Estimation of $\boldsymbol{\eta}$ may be obtained by the **discrete (or LC) MML approach** (Bartolucci, 2007)
- $\ell(\boldsymbol{\eta})$ may be efficiently maximize by the **EM algorithm** (Dempster et al., 1977)
- The software for the model estimation has been implemented in MATLAB and it can be downloaded from the web page:

  http://www.stat.unipg.it/bartolucci

Allocation to latent classes

For each specific item response pattern $\mathbf{x}$ it is possible to estimate the **posterior probabilities** of belonging to latent class c as follows

$$\hat{p}(\zeta_c|\mathbf{x}) = \hat{p}(\mathbf{\Theta} = \zeta_c|\mathbf{X} = \mathbf{x}) = \frac{\hat{p}(\mathbf{x}|\zeta_c)\hat{\pi}_c}{\sum_{h=1}^{C} \hat{p}(\mathbf{x}|\zeta_h)\hat{\pi}_h}, \quad c = 1, \ldots, C.$$

An individual is assigned to the latent class with the highest posterior probability.

Model selection

The specification of multidimensional LC Rasch model univocally depends on

- the number of latent classes C;
- the number of dimensions D and how items are associated to the different dimensions (through the dummies δ_{id}).

Choice of the number of latent classes

We base the selection of the optimal number of latent classes on the **Bayesian Information Criterion (BIC index)** of Schwarz (1978):

$$BIC = -2\hat{\ell} + g\ln(n),$$

where g is the number of free parameters and

$$g = (C-1) + DC + (k-D).$$

Considering that some parameters are constrained to 0 to ensure the model identifiability, there are:

- $C-1$ mass probabilities for the classes
- DC ability parameters
- $k-D$ difficulty parameters

Measurement of the correlation

For two dimensions d_1 and d_2 the correlation may be measured through the index

$$\hat{\rho}_{d_1 d_2} = \sum_{c=1}^{C} \hat{\zeta}^*_{cd_1} \hat{\zeta}^*_{cd_2} \hat{\pi}_c,$$

where $\hat{\zeta}^*_{cd}$, $c = 1, \ldots, C$, $d = 1, \ldots, D$, is the standardised estimate of the latent trait level referred to dimension d for the subjects in latent class c

Dimensionality analysis

- The hypothesis that two dimensions d_1 and d_2 are perfectly correlated is strongly related to the hypothesis that the items in $\mathcal{I}_{d_1}$ and $\mathcal{I}_{d_2}$ measure the same latent trait, i.e.

$$H_0 : \zeta_{cd_2} = \zeta_{cd_1} + a,$$

  where a is an arbitrary constant.

- We compare two main approaches:
  - LR test based on discrete MML estimates
  - LR test based on CML estimates or Martin-Löf test

Test of dimensionality based on discrete MML estimates

- The LR test statistic is given by

$$LR_1 = -2(\hat{\ell}_0 - \hat{\ell}_1)$$

- $\hat{\ell}_0$ is the maximum log-likelihood of the restricted (i.e. unidimensional) model
- $\hat{\ell}_1$ is the maximum log-likelihood of the general (i.e. bidimensional) model
- Under H_0, LR_1 is asymptotically distributed as a χ^2_{C-1}
- Main disadvantage of test based on LR_1 statistic:
  the results may depend on the choice of the number of latent classes

Martin-Löf test (1)

- Main references: Martin-Löf (1970), Glas and Verhelst (1995)
- The LR test statistic is given by

$$LR_2 = -2(\tilde{\ell}_0 - \tilde{\ell}_1)$$

- $\tilde{\ell}_0$ is the maximum log-likelihood of the restricted (i.e. unidimensional) model
- $\tilde{\ell}_0$ is obtained as the sum of the maximum conditional log-likelihood ($\tilde{\ell}_{0c}$) for the items in $\mathcal{I}_{d_1} \cup \mathcal{I}_{d_2}$ and the maximum marginal log-likelihood of the multinomial model for the distribution of the scores ($\tilde{\ell}_{0m}$), i.e.

$$\tilde{\ell}_0 = \tilde{\ell}_{0c} + \tilde{\ell}_{0m}$$

Martin-Löf test(2)

- $\tilde{\ell}_1$ is the maximum log-likelihood of the general (i.e. bidimensional) model
- $\tilde{\ell}_1$ is obtained as the sum of the maximum conditional log-likelihood ($\tilde{\ell}_{1c}^{(1)}$) for the items in $\mathcal{I}_{d_1}$, the maximum conditional log-likelihood ($\tilde{\ell}_{1c}^{(2)}$) for the items in $\mathcal{I}_{d_2}$, and the maximum marginal log-likelihood of the multinomial model for the distribution of the scores ($\tilde{\ell}_{1m}$), i.e.

$$\tilde{\ell}_1 = \tilde{\ell}_{1c}^{(1)} + \tilde{\ell}_{1c}^{(2)} + \tilde{\ell}_{1m}$$

- Under H_0, LR_2 is asymptotically distributed as a $\chi^2_{k_{d_1} k_{d_2} - 1}$
- Main disadvantages of Martin-Löf test:
  - its power is significantly affected by the length of the questionnaire and it may be disappointingly low with strongly correlated latent dimensions
  - it can only be used when Rasch paradigm holds

Rasch paradigm

- In order to detect dimensions coherent with the Rasch paradigm, we fit a separate Rasch model for each dimension and those that violate this paradigm are not included in the following analysis.
- The testing procedures (Linacre and Wright, 1994; Glas and Verhelst, 1995) we use are mainly based on the comparison with the two-parameter logistic model and with the saturated model by LR statistics, and on infit and outfit statistics.
- We select 5 dimensions:
  - 3 are referred to HrQoL (SF-36 questionnaire): Bodily Pain (BP), Social Functioning (SF), and Vitality (VT)
  - 2 are referred to psychopathological disturbs (HADS questionnaire): Anxiety and Depression

Discreteness assumption

In order to verify the discreteness assumption,

- we fit a separate (unidimensional) LC Rasch model for each dimension and for increasing values of the number of latent classes
- we fit a separate (unidimensional) normal Rasch model for each dimension
- we compare the two types of model in terms of maximum log-likelihood and BIC index

We observe that for each dimension, the unidimensional LC Rasch model always attains a higher maximum log-likelihood than its normal counterpart, while the minimum BIC index is always smaller.

Choice of the number of latent classes

Table 1: Maximum log-likelihood and BIC index for the multidimensional latent class Rasch model with a number of latent classes between 1 and 7; in bold are the data referred to the model with the smallest BIC.

	Number of latent classes (C)						
	1	2	3	4	5	6	7
$\hat{\ell}$	-2321.8	-2066.7	-2016.3	-1989.8	-1966.9	**-1949.3**	-1937.7
BIC	4760.6	4282.1	4213.2	4192.1	4178.3	**4175.0**	4183.7

Interpretation of latent classes (1)

Table 2: Estimated support points and probabilities of the latent classes (for each dimension, in bold the highest support point, in italic the smallest one).

	Latent class					
Dimension	1	2	3	4	5	6
BP	$-\infty$	*-2.17*	-0.92	-0.91	-0.20	**1.48**
SF	-2.56	$-\infty$	-1.71	-2.92	$-\infty$	**0.79**
VT	-2.36	*-3.28*	-0.46	**1.09**	-2.23	0.93
Anxiety	**0.00**	*-4.87*	-3.40	-1.85	-1.55	-0.05
Depression	**1.30**	*-4.10*	-1.99	0.12	-1.50	0.76
Probability	0.042	0.223	0.383	0.134	0.131	0.086

Interpretation of latent classes (2)

- Class 1
  - this is the smallest class
  - patients with the highest tendency to Anxiety and Depression
- Class 2
  - patients with the best conditions with respect to all dimensions
- Class 3
  - this is the largest class
  - patients with intermediate conditions with respect to each dimension
- Class 4
  - patients with the worst conditions with respect to VT
- Class 5
  - patients with the best conditions with respect to SF
- Class 6
  - patients with the highest impairments with respect to BP and SF

Correlation coefficients

Table 3: Estimated correlation matrix between latent traits.

	SF	VT	Anxiety	Depression
BP	0.679	0.794	**0.932**	**0.989**
SF		**0.934**	0.388	0.596
VT			0.530	0.707
Anxiety				**0.971**

Tests of dimensionality

Table 4: Results from the test of dimensionality based on the discrete MML parameter estimates; in brackets the p-values.

	SF	VT	Anxiety	Depression
BP	27.1 (0.000)	14.0 (0.007)	22.2 (0.000)	27.5 (0.000)
SF		12.5 (**0.014**)	32.1 (0.000)	26.3 (0.000)
VT			51.2 (0.000)	70.3 (0.000)
Anxiety				9.1 (**0.059**)

Table 5: Results from the test of dimensionality based on the CML parameter estimates (ML test); in brackets the p-values.

	SF	VT	Anxiety	Depression
BP	23.0 (0.000)	38.6 (0.000)	53.7 (0.000)	57.0 (0.000)
SF		30.7 (0.000)	42.9 (0.000)	36.3 (0.001)
VT			126.7 (0.000)	112.7 (0.000)
Anxiety				83.3 (0.001)

Conclusions

- The values of the test statistic based on discrete MML method are almost always smaller than the corresponding values computed on the basis of the CML method.
- Main conclusions are similar: both kinds of test do not show any evidence of unidimensionality.
- A weak evidence of unidimensionality only appears under the test based on the discrete MML estimates for anxiety and depression (p-value equal to 0.059) and for VT and SF (p-value equal to 0.014)

Therefore, the present study confirms that the analysed dimensions defined by the SF-36 and HADS questionnaires are indeed separate, and collapsing some of them would imply an inappropriate simplification of the phenomena investigated by the test items.

Further work

Extension of multidimensional LC IRT models

- in presence of ordinal polytomous items;
- to include covariates and to take into account Differential Item Functioning.

Main references

ADAMS R.J., WILSON M., WANG W.C., The multidimensional random coefficients multinomial logit model, *Applied Psychological Measurement*, vol. 21, 1997, 1-23.

BARTOLUCCI F., A class of multidimensional IRT models for testing unidimensionality and clustering items, *Psychometrika*, vol. 72, 2007, p. 141-157.

DEMPSTER A.P., LAIRD N.M., RUBIN D.B., Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Series B*, vol. 39, 1977, p. 1-38.

FAYERS P.M., MACHIN D., *Quality of life. Assessment, analysis and interpretation*, Wiley, 2000.

GLAS C.A.W., VERHELST N.D. Testing the Rasch model, in G.H. Fischer, and I.W. Molenaar (Eds.), *Rasch models: Foundations, recent developments, and applications*, New York, Springer-Verlag, 1995, p. 69-95.

GOODMAN L.A., Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, vol. 61, 1974, p. 215-231.

LAZARSFELD P.F., HENRY N.W., *Latent structure analysis*, Boston, Houghton Mifflin, 1968.

LINACRE, J. M., and B. D. WRIGHT. 1994. (Dichotomous mean-square) infit and outfit chi-square fit statistics. Rasch Measurement Transactions 8(2): 360.

Main references

MARTIN LÖF P., Statistical models: Notes from seminars in the academic year $1969 - 1970$, with the assistance of Rolf Sundberg. Stockholm: Institütet för Försäkringsmatemetik och Matematiks Statistiks vid Stockholms Universitet, 1970.

RASCH G., On general laws and the meaning of measurement in psychology, *Proceedings of the IV Berkeley Symposium on Mathematical Statistics and Probability*, vol. 4, 1961, p. 321-333.

SCHWARZ G., Estimating the dimension of a model. *Annals of Statistics*, vol. 6, 1978, p. 461-464.

WARE J.E., KOSINSKI M., DEWEY J.E., *How to Score Version 2 of the SF-36 Health Survey (Standard & acute forms)*. Lincoln RI, Quality Metric Incorporated, 2002.

ZIGMOND A.S., SNAITH R.P., The hospital anxiety and depression scale, *Acta Psychiatrika Scandinavica*, vol. 67 no. 6, 1983 p. 361-370.