Chapter 1

# Analysis of longitudinal HrQoL using latent regression in the context of Rasch modelling

## 1.1. Introduction

For the last ten years at least, both in the clinical tests and in the epidemiological studies the Health related Quality of Life (HrQoL) is considered a very important element to evaluate the state of health from the humane point of view and, therefore, to define the most suitable treatment. Particularly, in the context of palliative care for terminal cancer patients the HrQoL is not anymore a secondary end point of the treatment (being the survival the primary end point), but it is the primary end point. Therefore, the detection of good measurement and analysis tools is important to put information at clinicians' disposal to facilitate the therapy decisional process.

From a methodological point of view, decisions on two different but connected aspects have to be taken for the study of a latent variable, such as the HrQoL. The first one is about the definition of a suitable method of measurement to translate the qualitative information coming from a set of observable items, partial indicators of the latent variable, in a quantitative information. The second one is about the detection of a suitable statistical methodology to explain the latent variable.

The analysis presented in this chapter is developed in the framework of the Rasch-family models ([RAS 60]; [FIS 95b]) as regards the former aspect and in the framework of multilevel models for repeated measures [SNI 99] as regards the latter aspect.

Chapter written by Silvia BACCI, Department of Statistics "G. Parenti", University of Florence, Italy.

In the most common approach, the latent variable is estimated by means of a measurement model (for instance, a Rasch model) and *then* the estimates are used in a regression model as observed values of the response variable. However, some elements suggest caution toward a two separate steps approach: the bias and the inconstistency of the latent variable estimates ([GOL 80]; [LOR 84]), the underestimation of the true association between latent variable and covariates [MES 04] and, more in general, the lack of flexibility and integration between applied psychometricians and statisticians [WIL 04]. A possible solution is represented by a global approach that integrates the measurement phase and the analysis phase in the same model: in this case one talks of latent regression [AND 77]. The linear latent regression Rasch model is already known in literature since several years [ZWI 91], but the extension to longitudinal data has a more recent interest.

In this chapter, a longitudinal latent regression model that uses the Rasch analysis as measurement tool is proposed for the first time: it consists in a random intercept and slopes logistic model with covariates at the second aggregation level. The implementation of the model is performed in SAS, through the Nlmixed procedure: in such a way, it is possible to achieve a better flexibility for the utilization by users of general statistical softwares.

In section 1.2 a brief review about global models for longitudinal data is presented, whereas in section 1.3 the latent regression model for repeated measures is developed: in particular, the attention is focused on the model structure, on the correlation matrix, on the estimation and on the computational implementation. In section 1.4 the results of a case study on the HrQoL of terminal cancer patients under palliative care are shown. The applied analysis is specially focused on the time effect and on the significativity of baseline condition and of other variables on the HrQoL. Finally, some concluding remarks are reported.

### 1.2. Global models for longitudinal data analysis

Submitting the same questionnaire to the same set of people repetitively over the time is a rather common situation: the objective is monitoring the trend of a latent variable to understand if significant changes occur; then, it raises the problem to detect variables that explain these changes. In a context of this kind, the data structure is formed by observations for every person on more than one item, and, for every item, on more than one time point. In the global approach framework, methodological developments are concerned with three different types of models, though case studies are rather infrequent and related to rather simple examples: Linear Logistic model with Relaxed Assumptions (LLRA), Multi-dimensional Rasch model, Three-level regression model.

– *Linear Logistic Test Model with Relaxed Assumption* (LLRA)
Fischer [FIS 95a] formalized the measurement of change by means of the concept of

"virtual" items. Any change of the latent variable occurring between two time points can be described for every person as a change of the item parameters. An item $I_j$ given to the same person at two different time points, $T_1$ and $T_2$, can be considered as a pair of "virtual" items, $I_a^*$ and $I_b^*$, with associated a pair of "virtual" difficulty parameters, $\beta_a^*$ e $\beta_b^*$: if the amount of change between $T_1$ and $T_2$ is equal to a constant $\delta$ (over all people), then the pair of "virtual" items generated by real item $I_j$ with difficulty $\beta_j$ is characterized by the two parameters: $\beta_a^* = \beta_j$ and $\beta_b^* = \beta_j + \delta$. Hence, the item parameters are a linear combination of the real item parameter and the change effect. This kind of model belongs to Linear Logistic Test Model (LLTM) family. Substituting $(\theta_i - \beta_j)$ with only one parameter $\theta_{ij}$ a generalization of LLTM is obtained: it is named Linear Logistic model with relaxed assumptions and it is a LLTM with two (in the case of two repeated measures) items for every person.

   – *Multi-dimensional Rasch model*
Wang and Chyi-In [WAN 04] model the latent variable change over the time through a multi-dimensional Rasch model. They discern an initial ability parameter and a modifiability parameter for each occasion following the first one: all of these parameters are considered as different latent dimensions which describe the change among adjacent time points.

   – *Three-level regression model*
Pastor and Beretvas [PAS 06] start with the Rasch model formulation as two-level model, where the item responses are the first-level units and individuals are the second-level units. Then, they put beside this model a longitudinal regression model: it is also a two-level model, where measurement occasions are the first-level units and individuals are still the second-level units. Putting together the two models, a three-level model results with item responses aggregated in measurement occasions and measurement occasions aggregated in individuals. The output is a model that explains the change, where what is changing is the latent variable and, under every second-level unit, there is a measurement model.

   Among the three mentioned models, LLRA and multi-dimensional Rasch model are not very suitable for complex data structure. Indeed, for both of them the number of parameters increases by increasing the number of measurement occasions for every person. In detail:

   – LLRA: For each real item there are as many virtual items as measurement occasions; moreover, every virtual parameter is a linear combination of the real item parameter and of a change effect. Hence, if change effect is constant for all items, the number of parameters is equal to the number of measurement occasions; otherwise, if the change effect varies among items, the number of parameters is equal to the product of the number of measurement occasions and the number of real items.

   – Multi-dimensional Rasch model: The latent ability for every measurement occasion is a dimension; moreover, for every pair of dimension three parameters have to be estimated, i.e. two variances and one correlation coefficient. Therefore, every

time point added to the model causes new parameters: a variance and a correlation coefficient between the new dimension and each of the previous ones.

The multilevel model is, then, the most flexible, though the three-level structure is computationally heavy for the estimation process. Moreover, though the model does not become more complicated when the measurement occasions increase and it allows that different people have a different number of time points, a too small number of observations for every person (that is a too small number of second-level units) may cause inaccurate estimates and high standard errors.

### 1.3. A latent regression Rasch model for longitudinal data analysis

#### 1.3.1. *Model structure*

To remedy the above stated problems, an alternative model is proposed to analyze longitudinal data in the context of Rasch measurement models. The framework of reference is the latent regression: the aim is modelling the latent variable $\theta$ as a function of the time and of other covariates. Given the multilevel structure of longitudinal data (measurement occasions are first-level units and individuals are second level units), a (latent) regression model with random intercept is estimated to take into account the variability among individuals. Moreover, it is reasonable to assume that also the time effect - in the clinical context it is often the effect of a therapy - is variable among people. In the end, the dependent variable, i.e. the latent ability, is assumed continuously (in particular, normally) distributed.

Let us denote with: $i = 1, 2, \ldots, N$ the individuals; $t$ the measurement occasion; $\theta_{it}$ the "ability" parameter (i.e. the HrQoL level) for $i$-th person at $t$ time; $f(t)$ any function of $t$ (e.g. a linear or a quadratic function); $z_i$ the value of covariate $z$ assumed by the $i$-th person; $\alpha$ the regression coefficient of $z_i$; $\delta_{0i}$ and $\delta_{1i}$ the random intercept and random coefficient, respectively; $\gamma_{00}$ and $\gamma_{11}$ the fixed components of $\delta_{0i}$ and $\delta_{1i}$, respectively; $u_{0i}$ and $u_{1i}$ the random components of $\delta_{0i}$ and $\delta_{1i}$, respectively. More in detail, the random intercept $u_{0i}$ explains how the initial level (i.e. for $t = 0$) of $\theta_{it}$ of each patient differs from the average population value; instead, the random coefficient $u_{1i}$ means how the individual time effect differs from the average population time effect. A linear multilevel model is described by the following equations:

$$
\begin{cases}
\theta_{it} &= \delta_{0i} + \delta_{1i} \cdot f(t) + \alpha \cdot z_i \\
\delta_{0i} &= \gamma_{00} + u_{0i} \\
\delta_{1i} &= \gamma_{11} + u_{1i}
\end{cases}
$$

$$
\Updownarrow
$$

$$
\theta_{it} = (\gamma_{00} + \gamma_{11} \cdot f(t) + \alpha \cdot z_i) + (u_{0i} + u_{1i} \cdot f(t)) \tag{1.1}
$$

With reference to the second member of equation (1.1), the terms in the first brackets define the fixed part of the model, that is the mean $\mu_\theta$ of $\theta_{it}$, whereas the terms in the second brackets define the random part of the model. The structure of the random effects is:

$$\left[ \begin{array}{c} u_{0i} \\ u_{1i} \end{array} \right] \sim Normal \left[ \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \left( \begin{array}{cc} \sigma_{u0}^2 & \sigma_{u01} \\ \sigma_{u01} & \sigma_{u1}^2 \end{array} \right) \right] \tag{1.2}$$

With respect to other multilevel linear models, this is different because the response variable is not observed: hence, it has to be estimated by means of a Rasch model. For dichotomous items, the Rasch model is given by the following equation [FIS 95b], where the subscript $t$ is introduced here for coherence with eq. (1.1):

$$P(X_{jti} = 1|\theta_{it}; \beta_j) = \frac{exp[\theta_{it} - \beta_j]}{1 + exp[\theta_{it} - \beta_j]}$$

$$\Updownarrow$$

$$logit[P(X_{jti} = 1|\theta_{it}; \beta_j)] = \log \left[ \frac{P(X_{jti}=1|\theta_{it}, \beta_j)}{1 - P(X_{jti}=1|\theta_{it}, \beta_j)} \right] = \theta_{it} - \beta_j, \tag{1.3}$$

where: $i$ is the $i$-th person ($i = 1, 2, \ldots, N$); $j$ is the $j$-th item ($j = 1, 2, \ldots, J$); $x_{jti}$ is the response category of the $j$-th item, i.e. 0 or 1, chosen by the $i$-th person at $t$-th time; $\beta_j$ is the difficulty parameter of $j$-th item.

If $\theta_{it}$ is substituted in the Rasch model with the structural model of eq. (1.1), the following **longitudinal latent regression model** (**LLRM**) is obtained:

$$logit[P(X_{jti} = 1|z_i, u_{0i}, u_{1i}; \gamma_{11}, \alpha, \Sigma, \beta_j)] =$$
$$= (\gamma_{00} + \gamma_{11} \cdot f(t) + \alpha \cdot z_i - \beta_j) + (u_{0i} + u_{1i} \cdot f(t)), \tag{1.4}$$

where $\Sigma$ is the variance and covariances matrix of latent variable. In order that the model is identifiable, the intercept $\gamma_{00}$ of structural model has to be fixed at 0; alternatively, one of the difficulty parameters may be fixed and so $\gamma_{00}$ is free.

The proposed model is still a random intercept and slope model, but it is not linear, but logistic, and it allows a direct estimation of parameters from observed item responses. Several generalizations are possible. First of all, more than one covariate can be added and the computational heaviness does not increase substantially. Similarly, for polytomous items, the Partial Credit Model[1] [MAS 82] is simple to use by means of substituting eq. (1.1) in eq. (1.5), that is substituting $\beta_j$ with $\beta_{jk}$ parameters. Moreover, the model can be used to estimate Differential Item Functioning (DIF)

_____

1. Here is reminded the formula of Partial Credit Model:

effect on one or more items, that describes a different functioning of items among different groups of individuals. It is sufficient adding an interaction effect to item suspected of DIF: i.e. $\beta_j$ has to be substituted by $\beta_j \cdot z_{0i}$, where $z_{0i}$ is a dummy variable that indicates the belonging of $i$-th person to a group of people (e.g. males vs females). Moreover, by substituting $\beta_j$ with a random coefficient $\beta_{ji}$, one can test the hypothesis that a 2-parameter logistic model [BAK 04] describes better the data: $\beta_{ji}$ has got a fixed component, that means the average difficulty of item for the population, and a random component, that means how different the difficulty of item is for the $i$-th person. Still, if an interaction effect between a difficulty parameter $\beta_j$ and the time variable is added, the time stability of questionnaire can be evaluated: it is important to remind that, in order that a questionnaire is valid, it must be stable, that is each item must give a contribution to measure the same latent variable in every time point.

Output from LLRM model is similar of output from "classical" longitudinal models. If the main purpose of analysis is an evaluation of the time effect on the latent variable, the attention is on second level residuals estimates (the $u_{1i}s$). The $u_{1i}s$ show how time effect on latent variable for a specific person diverges from the average effect for all population. For example, in a clinical study, monitoring the impact of a therapy on patients could be interesting: so, the time effect describes the effect of the therapy (if the questionnaire is filled in repetitively during the therapy) and an $u_{1i}s$ residuals analysis allows to detect individual characteristics of patients with specially positive or negative reaction. This kind of information can be useful to produce a classification of patients based on individual characteristics to anticipate (before the therapy beginning) the impact that the therapy will have on the HrQoL of every patient.

### 1.3.2. *Correlation structure*

To facilitate the analysis and the understanding of the longitudinal latent regression model, it is useful to know the correlation structure of the latent variable $\theta_{it}$. By recalling eq. (1.1), the variances and covariances matrix elements are promptly defined:

$$P(X_{itj} = x_{ij}|\theta_{it}, \beta_{jx}) = \frac{exp[\sum_{k=0}^{x_{ij}}(\theta_{it} - \beta_{jk})]}{\sum_{h=0}^{H_j} exp \sum_{k=0}^{h}(\theta_{it} - \beta_{jk})} \tag{1.5}$$

where $x_{ij} = 0, 1, \ldots, h, \ldots, H_j$ and $\beta_{jk}$ means the difficulty level of the $k$-th threshold of the $j$-th item.

$$Var(\theta_{it}|z_i) =$$

$$= Var[\gamma_{00} + \gamma_{11} \cdot f(t) + \alpha \cdot z_i) + (u_{0i} + u_{1i} \cdot f(t)] =$$
$$= Var[u_{0i} + u_{1i} \cdot f(t)] =$$
$$= \sigma_{u0}^2 + f^2(t) \cdot \sigma_{u1}^2 + 2 \cdot f(t) \cdot \sigma_{u01};$$

$$Cov(\theta_{it}, \theta_{i't'}|z_i) =$$

$$= Cov[(u_{0i} + u_{1i} \cdot f(t); (u_{0i'} + u_{1i'} \cdot f(t')] =$$
$$= \{\sigma_{u0}^2 + f(t)f(t')\sigma_{u1}^2 + \sigma_{u01}[f(t) + f(t')]\} \cdot I_{(i=i')}.$$

These formulas show that the variance and covariances (and, therefore, the correlations) of the latent variable depend only on the measurement occasion: this means that the correlation between the measurements of the $i$-th person's latent variable in two different time points changes with the change of the time lag and, in particular, it decreases when the time lag increases (as it is shown in the applied analysis, Table 1.2). Naturally, coherently with the assumptions of the model, the correlation between the measurements of the latent variable for two different individuals is equal to $0$.

### 1.3.3. *Estimation*

There are three main approaches to estimate a Rasch model [BAK 04]: joint maximum likelihood (JML), conditional maximum likelihood (CML) and marginal maximum likelihood (MML). In JML and CML person parameters are considered fixed effects, whereas in MML they are assumed random and independent drawings from a density distribution that describes the population. The latent regression point of view considers the latent ability as a random variable: so, the most suitable approach is the MML. The log-likelihood function is obtained integrating out the conditional log-likelihood on the random variable; then, the log-likelihood for the longitudinal latent regression model is the following one:

$$\log L(\boldsymbol{\beta}, \gamma_{11}, \alpha, \boldsymbol{\Sigma}) = \sum_{i=1}^{N} \log \int_{\Theta} P(\mathbf{x_{it}}|\theta_{it}; \beta) \cdot \phi(\theta_{it}|\gamma_{11}, \alpha, \boldsymbol{\Sigma}) d\theta_{it}, \qquad (1.6)$$

where $\mathbf{x_{it}}$ is the response pattern for the $i$-th person at $t$ time, $\boldsymbol{\beta}$ is the difficulty parameters vector and $P(\mathbf{x_i}|\theta_{it}; \boldsymbol{\beta})$ is the probability of $\mathbf{x_{it}}$:

$$P(\mathbf{x_{it}}|\theta_{it}; \boldsymbol{\beta}) = \prod_{j=1}^{J} \frac{exp[x_{jti}(\theta_{it} - \beta_j)]}{1 + exp(\theta_{it} - \beta_j)}.$$

Because the integral in the marginal log-likelihood does not have a closed solution, a numerical integration method has to be used: in this chapter the adaptive Gaussian quadrature was adopted. Finally, the estimation of random effects is usually based on the posterior distribution of the latent variable: the mean or the mode of this distribution is used as point estimates of $u_{0i}$ and $u_{1i}$ (so-called empirical Bayes estimates).

### 1.3.4. *Implementation with SAS*

One of the main problems about the utilization of measurement models by common-users is that, on one side traditional computational tools do not have specialized routines to estimate this kind of models, and on the other side devoted softwares are hardly configurable. To remedy these problems, the most recent literature is addressed to study the potentialities of generical statistical softwares, *in primis* SAS (for instance, [DOR 03]; [WIL 04]; [HAR 07]) and Stata [RAB 04]. So, it is possible the estimation of Rasch models without using compulsorily devoted softwares, and the estimation of generalized Rasch models, such as latent regression models, that otherwise are not applicable. This chapter gets into this current of study, intending to implement in SAS the longitudinal latent regression model proposed in the previous section.

The multilevel structure of Rasch model - item responses are the first-level units and persons are the second-level units - allows to use the Nlmixed procedure of SAS [WIL 04]. In comparison with a simple Rasch model, the longitudinal latent regression model is still a multilevel logistic model with two aggregation levels and with a random intercept, but also with a random coefficient to explain the time effect. The random component of the model is not any more $\theta$, but it is formed by the two random effects $u_0$ and $u_1$, that appear in the regression model for $\theta$.

For a dichotomous Rasch model, let time be the linear time effect (more in general, any kind of time effect can be introduced), let z be a covariate with fixed effect, let beta1, ..., betaJ be the difficulty item parameters, let alpha be the regression fixed coefficient of z, let Response be the response vector, let theta be the latent variable, let Individual be the second level unit, let I1, ..., IJ be the indicators of items, let s2u0, cu01, s2u1 be the first-level variance, the covariance, the second-level variance, respectively, and let u0 and u1 be the random effects, then the SAS code to estimate the LLRM is the following one:

```
proc nlmixed data = filename qpoints = number;
parms beta1 = num1 beta2 = num2 ... betaJ = numJ alpha=numAlpha s2u0 =
numS2u0 s2u1 = numS2u1 cu01 = numCu01;
meantheta = gamma11 · time + alpha · z;
theta = meantheta + u0 + u1 · time;
eta = theta - (beta1 · I1 + beta2 · I2 + ... + betaJ · IJ);
expeta = exp(eta);
p = expeta / (1 + expeta);
model Response ~ binomial (1, p);
random u0 u1 ~ normal ([0, 0], [s2u0, cu01, s2u1]) subject = Individual out =
residual;
run;
```

Data have to be organized in a matrix so that there is not one line for each person, but rather one line for each person-item-measurement occasion combination. Note that the number of measurement occasions can change for different people.

The model developed in this section can also be implemented in Stata by means of Gllamm routine [RAB 04]. The analysis in the next paragraph has been implemented with both SAS and Stata: outcomes are the same, but Stata's computational times are longer (about 45 minutes in SAS and 4 hours and 30 minutes in Stata).

## 1.4. Case study: longitudinal HrQoL of terminal cancer patients

In this section is presented an application of the longitudinal latent regression model. The analyzed data are supplied by an Italian multicentric study, called "Staging", and they concern 485 terminal cancer patients under domiciliary palliative care. A questionnaire for measuring HrQoL has been submitted to every patient both before the beginning of the treatment (at baseline) and, afterwards once a week until death or, in most fortunate cases, until the end of study period[2]. At baseline also several (about 40) individual characteristics have been surveyed about the patient, his/her family, his/her house, his/her disease, his/her clinical situation at the moment of the first visit, in the previous week, in the previous month and in the previous year. The questionnaire is called TIQ (Therapy Impact Questionnaire [TAM 92]): it was implemented in 1987 at Pain Therapy and Palliative Care division of National Cancer Institute in Milano (Italy), to measure the HrQoL in cancer patients. The questionnaire is composed by 36 items with four ordinal response categories ("not at all" = high HrQoL, "some", "a lot", "very much" = very low HrQoL). The following analysis is concerned with only a subset of 8 items describing the psychological component of HrQoL:

---

2. The number of compiled questionnaires is very variable for different people: minimum is 1 (only survey at baseline) and maximum is 32 (survey at baseline and 31 weeks under palliative care).

difficulty in performing usual free time activities (diffree), fatigue (fatigue), illness feeling (illness), sad or depressed feeling (sad), difficulty in concentrating or paying attention (difconce), nervousness feeling (nervousness), insecurity (insecurity) and confusion feeling (confusion).

As already mentioned in the introduction, when the disease is at a terminal phase, the most important end point of the pain therapy is not any more the survival, but the quality of life of the patient. The aim of the survey carried out by the "Staging" study is twofold:

– evaluating the time effect on HrQoL trend during the palliative care;

– understanding if the measurement of HrQoL at baseline gives a sufficient information to predict the HrQoL during the therapy or, on the contrary, the estimation at baseline has to be completed with other information on individual characteristics.

The latent regression model was estimated on 285 patients with at least 3 measurements on quality of life (plus the measurement at baseline). The selection model occurred in subsequent steps, taking into account a different number of random effects and performing several explorative analysis to select the significant covariates. Finally, the following model has been selected:

$$
\begin{aligned}
logit&[P(X_{jti} = 1 | u_{0i}, u_{1i}; \mu_\theta, \boldsymbol{\Sigma}, \boldsymbol{\beta})] = \\
&= (0.271 \cdot \sqrt{t} + 0.825 \cdot HrQoL_{base_i} + 0.546 \cdot \mathsf{difconce} - 0.308 \cdot \mathsf{confusion} + \\
&+ 1.997 \cdot \mathsf{diffree} + 1.077 \cdot \mathsf{illness} + 0.302 \cdot \mathsf{insecurity} + 2.487 \cdot \mathsf{fatigue} + \\
&+ 1.275 \cdot \mathsf{sad} + 0.632 \cdot \mathsf{nervousness}) + (u_{0i} + u_{1i} \cdot \sqrt{t})
\end{aligned}
$$

To rightly understand the estimated regression coefficients, the dichotomization applied for the item response categories has to be taken into account: 0 for the category "not at all", that is absence of symptom or disorder, 1 for the categories "some", "a lot" or "very much", that is presence of symptom or disorder. Hence, the probability of category 1 rather than 0 is as higher as lower is the level of HrQoL; *vice versa*, patients with a good HrQoL are inclined to choose category 0. Therefore, an high numerical value (both at baseline and after) is related with a low level of HrQoL.

For further details, Table 1.1 shows the estimated values of parameters with standard errors, p-values and 95% confidence intervals; Table 1.2 shows the correlation matrix (for the first 8 measurement occasions). The correlation between the HrQoL of the same patient in two different time points is always highly positive, though it decreases when the time lag increases.

The outcomes from the estimated model lead into several considerations. As regards the main objective of analysis (detecting determinants of HrQoL during palliative care) only quality of life at baseline is significant with a positive effect (low values

|  | Estimate | Stand.Error | p-value | Lower limit | Upper limit |
|---|---|---|---|---|---|
| difconce | -0.546 | 0.088 | $< .0001$ | -0.719 | -0.373 |
| confusion | 0.308 | 0.873 | 0.0005 | 0.137 | 0.480 |
| diffree | -1.997 | 0.096 | $< .0001$ | -2.187 | -1.808 |
| insecurity | -0.302 | 0.088 | $< .0001$ | -0.474 | -0.129 |
| illness | -1.077 | 0.090 | 0.0007 | -1.253 | -0.901 |
| nervousness | -0.632 | 0.088 | $< .0001$ | -0.805 | -0.458 |
| fatigue | -2.487 | 0.102 | $< .0001$ | -2.687 | -2.287 |
| sad | -1.275 | 0.091 | $< .0001$ | -1.454 | -1.097 |
| $\sqrt{t}$ | 0.271 | 0.044 | $< .0001$ | 0.186 | 0.357 |
| $HrQoL_{base}$ | 0.825 | 0.062 | $< .0001$ | 0.703 | 0.947 |
| $\sigma_{u0}^2$ | 0.971 | 0.138 | $< .0001$ | 0.699 | 1.242 |
| $\sigma_{u1}^2$ | 0.272 | 0.041 | $< .0001$ | 0.190 | 0.351 |
| $\sigma_{u01}$ | -0.139 | 0.060 | 0.0222 | -0.257 | -0.020 |

**Table 1.1.** *Longitudinal latent regression model: estimates, standard errors, p-values, confidence intervals ($\alpha = 0.05$).*

|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1.000 | 0.860 | 0.743 | 0.650 | 0.575 | 0.514 | 0.463 | 0.420 | 0.383 |
| 1 | | 1.000 | 0.978 | 0.945 | 0.910 | 0.878 | 0.849 | 0.823 | 0.800 |
| 2 | | | 1.000 | 0.992 | 0.975 | 0.956 | 0.937 | 0.919 | 0.903 |
| 3 | | | | 1.000 | 0.996 | 0.986 | 0.975 | 0.963 | 0.951 |
| 4 | | | | | 1.000 | 0.997 | 0.991 | 0.984 | 0.976 |
| 5 | | | | | | 1.000 | 0.998 | 0.994 | 0.989 |
| 6 | | | | | | | 1.000 | 0.999 | 0.996 |
| 7 | | | | | | | | 1.000 | 0.999 |
| 8 | | | | | | | | | 1.000 |

**Table 1.2.** *Longitudinal latent regression model: correlation matrix (first 8 measurement occasions).*

of HrQoL at baseline are related with low values of HrQoL during the treatment and *vice versa*). None of the 40 individual characteristics observed at baseline gives significant information on the future trend of HrQoL. Then, about the initial question on what variables have to be gathered at baseline to predict the HrQoL trend, the model shows that the questionnaire filled at the beginning by patient can explain some variability of future HrQoL; whereas the knowledge of further characteristics does not add any relevant information.

The other significant variable is time effect: time affects numerical value of HrQoL with a quadratic positive effect. In particular, the value of $0.271$ for the regression coefficient means that when the time lag increases the numerical value of HrQoL increases (i.e. its level gets worse), with a marginally decreasing rate. The deterioration of HrQoL is explainable with the approach to the death. The time effect can be considered as the result of two different, but hardly separable effects: approach to death and palliative care. Understanding if palliative care has a positive effect on the HrQoL should be an interesting question. Unfortunately, the data set does not give any aid to answer to this question: the study should be repeated with a control group, so that to compare patients under palliative care and patients under a different kind of therapy.

The fixed time coefficient equal to $0.271$ defines the *average* time effect on the population. Second-level residuals ($u_{1i}$), on the other hand, show how much time effect for the $i$-th person goes away from the average value of population. Positive values mean an individual time effect greater than average, that is an impact on quality of life more negative; negative values mean an individual effect less negative and, particularly, values smaller than $-0.271$ cause an improvement of HrQoL over the time. Figure 1.1 shows a classification of patients based on $u_{1i}$ residuals. For every patient the residual value and the corresponding confidence interval (based on suggestions of Goldstein and Healy [GOL 95]) are shown: two patients can be considered significantly different when the respective intervals are not overlapped. How the graphic shows, only individuals with extreme residuals are significantly different from the remaining population. In particular, only 15 patients (called Group 1) show a confidence interval with values smaller than $-0.271$: in other words, at a confidence level of 95%, the HrQoL of these people has an improvement over the time. On the opposite side, the confidence interval of 77 patients (called Group 2) contains only values greater than $-0.271$, so the individual time impact on HrQoL is negative at 95% level. Table 1.3 shows the individual characteristics that have a different distribution between the two groups of patients. From a descriptive point of view, patients from Group 1 are distinguished by a lower percentage of people that live alone and, consequentially, by an higher percentage of people that have a partner; more than $1/3$ of them has not any metastasis versus the 11% of patients from Group 2; moreover, some physical symptoms, such as break through pain, endocranial hypertension and decubitus lesions, are more rare. Finally, only 36% (versus 59%) are confined to bed and 60% (versus 74%) depend on someone else for the activity daily living. Really, the differences between the two groups of patients are statistically significant only in relation with the absence of metastasis (1% significativity level) and the compulsion to stay in bed (5% significativity level).

The other second-level residual component, $u_{0i}$, explains the alteration of intercept for average patient in $t = 0$, that is after the first week of therapy. The interpretation is similar to $u_{1i}s$: positive values mean an initial HrQoL level worse than average population and negative values mean a level better than average. The covariance between second level residuals is negative; the correlation coefficient between $u_{0i}$ and $u_{1i}$ is

|  | Group 1 | | Group 2 | |
| --- | --- | --- | --- | --- |
|  | % | Total | % | Total |
| lives alone | 0.07 | 1 | 0.16 | 12 |
| has a partner | 0.80 | 3 | 0.58 | 32 |
| no metastasis** | 0.36 | 5 | 0.11 | 8 |
| break through pain | 0.07 | 1 | 0.19 | 13 |
| endocranial hypertension | 0.00 | 0 | 0.08 | 6 |
| decubitus lesion | 0.00 | 0 | 0.10 | 8 |
| confined to bed* | 0.36 | 5 | 0.59 | 44 |
| dependent for adl | 0.60 | 9 | 0.74 | 57 |
| Total |  | 15 |  | 77 |

**Table 1.3.** *Comparison between patients with a positive individual time effect on HrQoL (Group 1) and patients with a negative individual time effect on HrQoL (Group 2). (\*\* = significativity at $1\%$ level; \* = significativity at $5\%$ level.)*

equal to $-0.270$. This means that patients with a HrQoL level better than average after the first visit ($t = 0$) have a worse reaction during the continuation of therapy; *vice versa*, the therapy should be more effective for people with a more crucial situation at beginning. A possible interpretation of negative correlation can be ascribed to the characteristics of palliative care, whose effectiveness level appears better in more crucial situations.
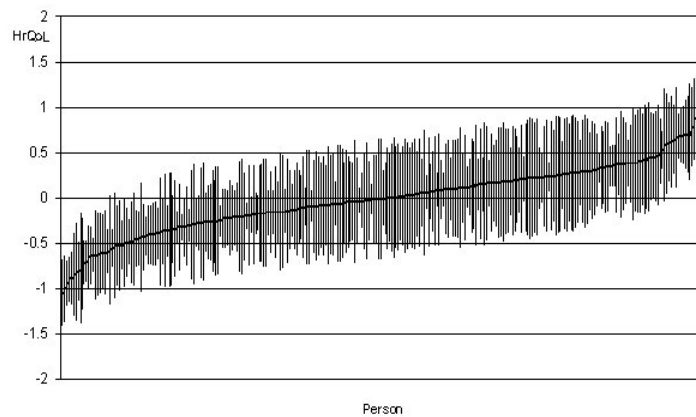


**Figure 1.1.** *Second-level residuals $u_{1i}$: estimates and confidence intervals.*

In the end, the estimation of the LLRM gives information about the difficulty of items: diffree and fatigue items are the "easiest", that is they are attended in the many of patients, also patients with a good HrQoL level; instead, confusion and insecurity items are the most "difficult", that is only people with a low HrQoL level show them.

To conclude the analysis, a two steps approach was applied, separating the measurement model from the structural model. First of all, the HrQoL of every person in different time points has been estimated; *then*, these values were used as dependent variable in a multilevel linear model (see eq. (1.1)), with the same covariates of LLRM (without items, obviously):

$$Y_i = (\gamma_{00} + \gamma_{11} \cdot \sqrt{t} + \alpha \cdot HrQoL_{base_i}) + (u_{0i} + u_{1i} \cdot \sqrt{t} + \epsilon_{ji}).$$

Outputs from the two steps approach are similar with outputs coming from other analysis ([ZWI 91] and [ADA 97]): as Table 1.4 shows, both time effect and HrQoL at baseline coefficient estimated by the two steps approach are smaller than the corresponding values estimated by the latent regression model.

|  | LLRRM | Long.Linear model |
| --- | --- | --- |
| $\sqrt{t}$ | 0.271 | 0.035 |
| $HrQoL_{base}$ | 0.825 | 0.542 |
| $\sigma_{u0}^2$ | 0.971 | 0.026 |
| $\sigma_{u1}^2$ | 0.271 | 0.006 |
| $\sigma_{u01}$ | -0.139 | -0.005 |

**Table 1.4.** *Global approach and two steps approach: comparison.*

## 1.5. Concluding remarks

This chapter is concerned with the relationship between measurement models and structural models when the interest of analysis is focused on a latent variable. A global approach that integrates measurement and explanation of a latent variable is more suitable than an approach that separates the two aspects. Following this idea, a latent regression model for longitudinal data is developed by using a Rasch model as measurement tool: it consists of a logistic model (or ordinal logistic in case of polytomous responses) with a random intercept and a random coefficient, which relates

explanatory covariates of latent variable *directly* with item responses. The model was implemented by means of the Nlmixed procedure of SAS.

The longitudinal latent regression model has got several advantages. First of all, it gets over the drawbacks of the two steps approach. Secondly, it is developed in a framework (multilevel modelling) well-known in the literature: this allows to estimate the model by means of generical statistical tools, supporting the utilization of latent regression by common-users. Moreover, it can be easily interpreted on the basis of random and fixed coefficients: in particular, the estimation of random components gives information on how each individual differs from the average population. In comparison with other global models for longitudinal data, its complexity does not increase with the number of measurement occasions; moreover, the treatment of a different number of observations for every individual does not represent any particular problem. Finally, as outlined in section 1.3.1, it can be extended to take into account more general data structures.

A priority question to study in deep in a following analysis is concerned with the belongings of the proposed model to the Rasch family. In other words, it is interesting to understand if the LLRM verifies the main properties of Rasch models, i.e. the sufficiency of the raw scores and the specific objectivity.

Moreover, the complexity of the analyzed data set illustrates some other problems that will be dealt in a future development of the work. First of all, the logistic latent regression model should be extended to consider also one or more latent covariates, so as to take into account their random nature, and to consider the multi-dimensional nature of the questionnaire. Secondly, another problem is the presence of informative drop out, due to the death of patients during the therapy.

# Bibliography

[ADA 97]  ADAMS R., WILSON M., WU M., "Multilevel item response models: an approach to errors in variables regression", *Journal of Educational and Behavioral Statistics*, vol. 22, num. 1, p. 47–76, 1997.

[AND 77]  ANDERSEN E., MADSEN M., "Estimating the parameters of the latent population distribution", *Psychometrika*, vol. 42, p. 357–374, 1977.

[BAK 04]  BAKER F., KIM S., *Item response theory. Parameter estimation techniques*, Dekker, 2004.

[DOR 03]  DORANGE C., CHWALOW, MESBAH M., "Analysing Quality of Life data with the ordinal Rasch model and NLMIXED SAS procedure", *Proceeding of the International Conference on Advance in Statistical Inferential Methods "ASIM 2003"*, 2003.

[FIS 95a]  FISCHER G., "Linear logistic models for changes", FISCHER G., MOLENAAR I., Eds., *Rasch models. Foundations, recent developments, and applications*, p. 157–180, Springer-Verlag: New York, 1995.

[FIS 95b]  FISCHER G., MOLENAAR I., *Rasch models. Foundations, Recent Developments and Applications*, Springer-Verlag, 1995.

[GOL 80]  GOLDSTEIN H., "Dimensionality, bias, independence and measurement scale problems in latent trait test score models", *British Journal of Mathematical and Statistical Psychology*, vol. 33, p. 234–260, 1980.

[GOL 95]  GOLDSTEIN H., HEALY M., "The graphical presentation of a colection of means", *Journal of the Royal Statistical Society series A*, vol. 158, p. 175–177, 1995.

[HAR 07]  HARDOUIN J., MESBAH M., "The SAS macro-program %ANAQOL to estimate the parameters of Item Response Theory models", In press on Communications in Statistics, 2007.

[LOR 84]  LORD F., *Maximum likelihood and bayesian parameter estiomation in IRT*, Educational Testing Service: Princeton, 1984.

[MAS 82]  MASTERS G., "A Rasch model for partial credit scoring", *Psychometrika*, vol. 47, p. 149–174, 1982.

[MES 04]  MESBAH M., "Measurement and analysis of health related quality of life and environmental data", *Environmetrics*, vol. 15, num. 5, p. 473–481, 2004.

[PAS 06]  PASTOR D., BERETVAS S., "An illustration of longitudinal Rasch modeling in the context of psychotherapy outcomes assessment", *Applied Psychological Measurement*, vol. 30, p. 100–120, 2006.

[RAB 04]  RABE-HESKETH S., SKRONDAL A., PICKLES A.,  "GLLAMM Manual", http://www.gllamm.orgh, 2004.

[RAS 60]  RASCH G., *Probabilistic models for some intelligence and attainment tests*, Danish Institute for Educational Research:Copenhagen, 1960.

[SNI 99]  SNIJDERS T., BOSKER R., *Multilevel analysis. An introduction to basic and advanced multilevel modeling*, Educational Testing Service: Princeton, 1999.

[TAM 92]  TAMBURINI M., ROSSO S., GAMBA A., MENCAGLIA E., DE CONNO F., VENTAFRIDDA V., "A therapy impact questionnaire for quality-of-life assessment in advanced cancer research", *Annals of Oncology*, vol. 3, p. 565–570, 1992.

[WAN 04]  WANG W., CHYI-IN W., "Gain score in item response theory as an effect size measure", *Educational and Psychological Measurement*, vol. 64(5), p. 758–780, 2004.

[WIL 04]  WILSON M., DE BOECK P., *Explanatory item response models: a generalized linear and nonlinear approach*, Springer-Verlag:New York, 2004.

[ZWI 91]  ZWINDERMANN A., "A generalized Rasch model for manifest predictor", *Psychometrika*, vol. 56, num. 4, p. 589–600, 1991.