Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal. Vol. 3, Issue 1 (2010), 44 – 51
ISSN 2070-5948, DOI 10.1285/i20705948v3n1p44

RASCH ANALYSIS AND MULTILEVEL MODELS FOR THE EVALUATION OF THE CUSTOMER SATISFACTION

Antonio Lucadamo*

Tedass, Università del Sannio, Benevento, Italy

Received 16 November 2009; Accepted 27 December 2009
Available online 28 February 2010

Abstract: *The evaluation of attitudes, capabilities and individual satisfaction is one of the most important problem of experimental sciences. These qualities in fact, are not directly observable, but they are expressed with polycotomous measure scales. To perform the evaluation it's necessary to substitute the qualitative modalities with some scores. These measures can be determined in different ways, but problems of quantification or relative to the conditions in which the survey is conducted can arise. For solving these problems one solution can be the Rasch Analysis. We use this technique to quantify the response to an hospital survey about the Customer Satisfaction. In a second step we try to verify if the patient satisfaction can be influenced by socio-economic factors and, for this reason, we use a Multilevel Model.*

***Keywords**: Rasch Analysis, Multilevel Models, Customer Satisfaction.*

1. Introduction

The evaluation of capabilities, attitudes, customer satisfaction, in different fields of the real life, is one of the problems that, during the last years, have been studied in social sciences. In fact, these qualities, are not directly observable, but they are expressed with the modalities of ordinal scales. In these circumstances, if we want to perform the evaluation, it is necessary to assign some scores to the categorical modalities of the analysis. The quantification is one of the most relevant problems in Customer Satisfaction. In fact, first of all it is necessary to keep in consideration the different distances between the modalities and, furthermore, there can be difficulties, due to the fact that there are often discrepancies between the value that a character

*Email: alucadam@unina.it

can assume when it is estimated in different times or situations or for different samples. To solve this kind of problems, one of the most used technique is the Rasch Analysis [1]. There are different Rasch models according to the characteristics that the items can assume [2]. This analysis is the start point of our paper. In fact, our purpose is to evaluate the Customer Satisfaction of the patients of an hospital. They participated to a survey that had the aim to judge the satisfaction related to different aspects of their staying in the hospital. The quantification of the response is made by the use of the Rasch Analysis and, in the second step of the work, we try to verify if the patient satisfaction can be influenced by socio-economic factors, by the qualities of the structure in which they are nursed and by the characteristics of doctors or nurses; for this reason we use a Multilevel Model.

The paper is organised as follows. Section 2 resumes the Rasch Analysis with special emphasis on the different kinds of models. Section 3 briefly describes the Multilevel Models. In Section 4, we apply some of the models previously described to data collected in an Italian hospital. Finally, conclusions are given in Section 5.

2. Rasch Analysis

Rasch models give the probability that an individual chooses a certain modality when he answers to a particular item [7]. This probability depends on different parameters as the ability of the individual, the difficulty of the item and the reached threshold. This result is obtained with the construction of probabilistic measurements derived by calibration methods, calculated considering unidimensional items. These kinds of measurements are linear, quantitative, test free and sample free [11]. This means that they are not related to the considered items and are not dependent on the characteristics of the people that are in the sample. This is possible because the Rasch model has three important properties: the additivity, the separability and the specific objectivity. The first one asserts that the measurements built with the Rasch Analysis are direct functions of the individual abilities and inverse functions of the difficulty of the item and of the difficulty to reach the thresholds. According to the second property, the measures don't depend on the sample and on the test difficulty. The last property instead affirms that the probabilities can be expressed in logit scores [8].

The simplest Rasch Model is the dichotomous model in which we consider only two possible options, "yes" or "not" that are generally codified as "1" and "0" [12]. The model can be expressed as follows:

$$\phi_{ni1} = \frac{\exp(\beta_n - \delta_{i1})}{1 + \exp(\beta_n - \delta_{i1})} \tag{1}$$

where ϕ_{ni1} is the person's probability of scoring 1 rather than 0 on item *i*, β_n is the ability of the individual *n* and δ_{i1} is the difficulty of the one step in item *i*.

The dichotomous model is the basis to build the model that have more than 2 modalities. They are the Partial Credit Model and the Rating Scale Model. In both the cases we have *m* modalities, where $m > 0$. In the Partial Credit, if we suppose that an individual *n* has reached the level $k-1$ in a defined item, the probability that the same person will reach the level *k* too is:

$$\phi_{nik} = \frac{\exp(\beta_n - \delta_{ik})}{1 + \exp(\beta_n - \delta_{ik})} \tag{2}$$

In this formula β_n is again the ability of the individual *n*, while δ_{ik} is the difficulty in reaching the level *k*. We can now calculate the probability that the individual *n* assign the value *k* in the item *i:*

$$\pi_{nik} = \frac{\exp \sum_{j=0}^{k} (\beta_n - \delta_{ij})}{1 + \exp \sum_{j=0}^{m} (\beta_n - \delta_{ij})} \qquad k = 1, \ldots, m \tag{3}$$

The numerator of the previous expression considers only the difficulties for the k passed steps. The denominator is, instead, the sum of all possible numerators. The difference between the Partial Credit and the Rating Scale is related to the difficulty that every person has in reaching a certain modality. In the Partial Credit Model in fact, every item can assume different difficulties in passing from step *k-1* to the step *k*. In the Rating Scale model instead, the parameter δ_{ik} is divided in two parts:

$$\delta_{ik} = \delta_i + \tau_k \tag{4}$$

δ_i is the scale value of the item and τ_k is the threshold for the modalities. In this case the threshold is the same over all the items. The probability that the individual *n* reaches the k level in the item *i* is:

$$\pi_{nik} = \frac{\exp \sum_{j=o}^{k} [\beta_n - (\delta_i + \tau_j)]}{1 + \exp \sum_{j=0}^{m} [\beta_n - (\delta_i + \tau_j)]} \tag{5}$$

Whatever is the model that we use, we need of an estimation criterion to obtain the parameter values. Different procedures have been developed, for example the PROX, the PAIR, the CON and the UCON procedure. In our analysis, we use the PROX procedure that will be described in section 4. In next paragraph we introduce instead the multilevel models.

3. Multilevel Models

Multilevel models suppose that in a hierarchical structure, the upper levels can influence the lower ones [10]; for example in an hospital the department can influence the satisfaction of the patients [9]. The basic model is an empty model defined as follows:

$$Y_{ng} = \gamma_{00} + U_{0g} + R_{ng} \tag{6}$$

In this formula, that can be viewed as the ANOVA with casual effects [3], there is a dependent variable Y_{ng} given by the sum of a general mean (γ_{00}), a random group effect (U_{0g}) and a random individual effect (R_{ng}). In this way the variability is divided in two parts [4]: in fact, in this model, it's assumed that the random variables U_{0g} and R_{ng} are mutually independent, normally distributed with mean equal to 0 and variances equal to τ^2 and σ^2. The total variance is then the sum of the two variances and we can compute the intra-class correlation coefficient:

$$\rho = \tau^2 / (\tau^2 + \sigma^2) \tag{7}$$

If this coefficient is significant, it is possible to carry out a Multilevel Analysis [5]. A first model is the Random Intercept Model that can be defined as follows:

$$Y_{ng} = \alpha_{0g} + \alpha_1 x_{ng} + R_{ng} \quad \text{with} \quad \alpha_{0g} = \gamma_{00} + U_{0g} \tag{8}$$

where x_{1g} are the variables related to the individuals. We can see that if the subscript g was not in the equation we had the classic linear regression model.
In the same equation if we consider the g subscript for the coefficient α_1 we will have the Random Slopes Model [6]:

$$Y_{ng} = \alpha_{0g} + \alpha_{1g} x_{ng} + R_{ng} \text{ with } \alpha_{0g} = \gamma_{00} + U_{0g} \text{ and } \alpha_{1g} = \gamma_{10} + U_{1g} \tag{9}$$

4. The Customer Satisfaction Analysis

The analysis concerns a survey effected in an hospital of the south of Italy. The patients answered to 30 questions about the services received during their staying in the hospital. For every question they could give a judgment between 1 and 7. Furthermore they furnished socio-economic information as, for example, the age, the gender, the income and the level of instruction. In the first step of the analysis we use the Rasch model, because it allows us to underline the individual attitude to give a positive judgment to the question. Furthermore, with this technique it is possible to detect the differences between the evaluation that every item has received. As we stated previously, we use, for our analysis, the PROX procedure. It assumes that the abilities of the individual and the scale value are more or less normally distributed; it is composed by different phases [7].

- First of all it is necessary to delete from the data-set the perfect scores: so all the individuals that assign value 1 or 7 to all the questions are not considered in the analysis. The same operation must be done on the items: so all the questions that received all the judgments equal to 1 or 7 must be deleted.
- The second operation is to linearize activity scores: this phase consists in transforming all the scores of a question (that is the sum of the values that the item received from all the

individuals) in a proportion P_i. This quantity is obtained by dividing the scores for their maximum values. This proportion is then transformed in a logit scale: $l_i = \log\left(\frac{1-P_i}{P_i}\right)$.

- The third operation is to center the activity logits, by subtracting their own mean ($d_i = l_i - \bar{l}_i$). This locates the origin of the logit scale at the mean of the activities.
- The last phase is necessary to remove the sample dispersion. If we are willing to work with the assumption that the sample can be satisfactorily described by a normal distribution, then the expansion factor needed to make this adjustment is $C_i = \sqrt{[((1+V)/2.89)/(1-UV/8.35)]}$ where V is the adjusted variance for the individuals and U is the corrected variance for the activities. This factor is multiplied for the centered scores and in this way we obtain the final calibration: $d_i^* = d_i C_i$. The linearization operations are repeated for the individuals too, considering in this case the scores obtained by every subject and taking into account the proportion, now indicated as P_n, and the logarithm $l_n = \log\left(\frac{P_n}{1-P_n}\right)$. To eliminate the sample dispersion the correction factor is $C_n = \sqrt{[((1+U)/2.89)/(1-UV/8.35)]}$ and the final estimation will be: $d_n^* = d_n C_n$ with $d_n = l_n - \bar{l}_n$.

Using this procedure we obtain the measure for the individuals and for the single item too. The measurements for the items that received the best scores are in table 1.

Table 1. Scores for some items

Items	**Estimated value**
Attention of the doctors on the problems	-0.66
Availability of the doctors	-0.55
Information about the disease	-0.44
Human relations	-0.42
Spaces in the waiting room	0.66
Cleanliness in the toilette	0.73
Quality of the food	0.85

They allow us to observe the aspects of the staying in the hospital that produce more satisfaction for the patients. An important consideration is about the scores of the items. As we can see in the table, the items with lower score are the ones that received the best judgments. In fact the score is computed, as we described, as the logarithm of the probability of not choosing that modality, divided by the probability of choosing it. The score is then the difficulty in obtaining high evaluation.

It’s easy to see that the items that had a major impact on the satisfaction are related to the medical area and to the human aspects. The structural aspects of the hospital received instead the worst judgments.

The same kind of results is computed for the individuals; in this case we obviously don't show the results in a table, but it's important to underline that the scores they obtained are considered as the overall satisfaction for every individual.
Once we obtain the value of the Customer Satisfaction, we want to verify if it could depends on some explicative variables. The linear regression model is not the right model for this kind of analysis. In fact, our data are organized in a hierarchical way. On the first level we have the patients and on the second level we can consider the departments of the hospital. On this second level, the explicative variable that we can consider is the experience of the doctors.
Before carrying out the analysis we standardize the variables. In this way we can better evaluate the impact of the explicative variables on the dependent one. In the second step of the analysis we verify the presence of an effect group and, afterward, we build the Random Intercept and the Random Slopes models. The second model doesn't show significant results, while the Random Intercept Model, described in equation 10, shows better results.

$$CS_{ng} = \alpha_{0g} + \alpha_1 (Age_{ng}) + \varepsilon_{ng} \qquad (10)$$

In the previous expression, the intercept can be written as $\alpha_{0g} = \gamma_{00} + \gamma_{01}(Exp_g) + R_{0g}$, the variable Age_{ng} is the age of the patients and Exp_g is the experience of the head physician of the departments, expressed in years. The estimated parameters are in table 2.

Table 2. Estimation of Random Intercept Model

Parameters	Coefficients	St. Error	T- ratio	P-value
γ_{00}	0,029342	0,052140	0,563	0,584
γ_{01}	0,391126	0,054041	7,238	0,000
α_1	0,191025	0,055735	3,427	0,001

As we can see the age of the patients has a positive coefficient and so, the older patients show a major level of satisfaction. A positive impact on the satisfaction is given also by the department in which they are. In fact if the ages of experience of the head physician increase, the value of the Customer Satisfaction will be higher. In figure 1 we can see how the Customer Satisfaction increases when the age of the patients and the experience of the doctors increase.

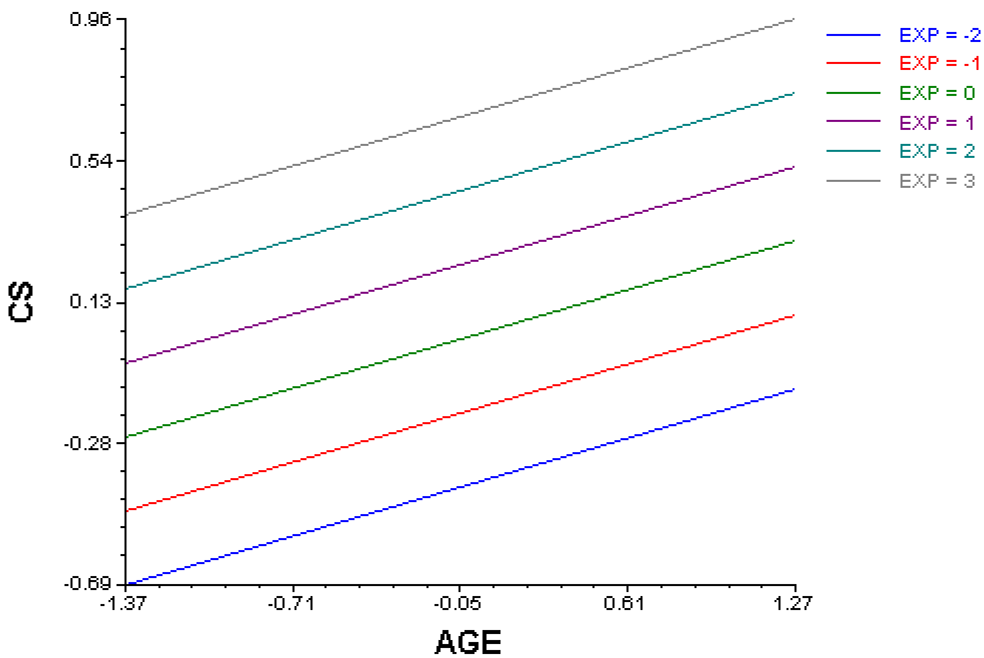

Figure 1. – Customer Satisfaction as function of the age of the patients and of experience of the head physician (standardized value)

5. Conclusions

In this paper we studied the Customer Satisfaction using two well known techniques. The Rasch analysis allowed us to obtain the scores about the satisfaction both for every patients and for the considered items. The scores obtained in the first phase are then been used as the basis for a multilevel analysis. This analysis showed the variables that had an higher impact on the patient satisfaction. For further work the analysis could be extended to more than a level or it could be possible to consider a cross section analysis. In this analysis the patients can be classified according both to the department and to the building, because departments and buildings are not always the same. A deeply study could instead be an analysis performed considering the group effects directly in the Rasch Analysis and not with the Multilevel Models. In this way we could have a Rasch Multilevel Model.

References

[1]. Bacci, S. (2006). I modelli di Rasch nella valutazione della didattica universitaria. *Statistica Applicata*, 18 (1), 1 – 45.

[2]. Bond, T.G., Fox, C.M. (2007). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Mahwah NJ: Lawrence Eribaum Associates.

[3]. Ciavolino, E., Dahlgaard, J.J. (2007). ECSI: Customer Satisfaction Modelling and Analysis: A Case Study. *Total Quality Management & Business Excellence*, 18 (5), 545 – 554.
[4]. Goldstein, H. (1995). *Multilevel statistical models*, London; Edward Arnold.
[5]. Hox, J.J. (1995). *Applied multilevel analysis*, Amsterdam: TT-Publications.
[6]. Hox, J.J. (2002). *Multilevel Analysis. Techniques and Applications*. Mahwah NJ: Lawrence Erlbaum Associates.
[7]. Linacre, J.M. (1989). *Many-Facet Rasch Measurement*, Chicago: MESA press Chicago.
[8]. Lunz, M.E., Wright, B.D. Linacre, J.M. (1990). Measuring the Impact of Judge Severity on Examination Scores. *Applied Measurement in Education*, 3 (4), 331 – 345.
[9]. Pagano, A, Rossi, C. (1999). La valutazione dei servizi sanitari in *Qualità e valutazione nei servizi di pubblica utilità*, eds E. Gori and G. Vittadini, ETAS Libri, Milano.
[10]. Snijders, T.A.B., Bosker, R.J. (1999). *Multilevel Analysis. An introduction to basic and advanced multilevel modeling*. London: SAGE Publications.
[11]. Wright, B.D. (1967). *Sample-Free test Calibration and Person Measurement*. Chicago: National Seminar on Adult Education Research.
[12]. Wright, B.D. and Masters, G.N. (1982). *Rating Scale Analysis. Rasch Measurement,* MESA press, Chicago.