DESCRIPTIVE ANALYSIS OF STUDENT RATINGS

Donata MARASINI [1]
PhD, Full Professor,
Department of Statistics ,
University of Milan – Bicocca, Italy

E-mail: donata.marasini@unimib.it

Piero QUATTO [2]
PhD, Associate Professor,
Department of Statistics ,
University of Milan – Bicocca, Italy

E-mail: piero.quatto@unimib.it

Abstract: *Let X be a statistical variable representing student ratings of University teaching. It is natural to assume for X an ordinal scale consisting of k categories (in ascending order of satisfaction). At first glance, student ratings can be summarized by a location index (such as the mode or the median of X) associated with a convenient measure of ordinal dispersion. For instance, the median of X may be associated with the dispersion index of Leti, resulting in a synthesis that takes into account the ordinal nature of data and also communicates information in an effective way. More generally, there are many indexes (such as the ordinal entropy) that can be properly employed to measure the ordinal dispersion. On the other hand, student ratings are often converted into scores and treated as a quantitative variable. More generally, it is possible to measure student satisfaction by means of a real-valued function defined on the standard simplex and satisfying some appropriate conditions. Finally, such a measure of satisfaction can be associated with a suitable measure of variability.*

Key words: *Teaching evaluation, Student ratings, Satisfaction measures, Ordinal dispersion*

1. Introduction

The evaluation of University teaching by students is generally carried out by administering a questionnaire. Such a questionnaire requires responses on an ordinal scale consisting of *k* categories $m_1 < \ldots < m_k$ in ascending order of satisfaction.

For example, the question "Overall, are you satisfied with this course?" (in the questionnaire adopted by Italian Universities) can be answered by "Decidedly no", "More no than yes", "More yes than no" or "Decidedly yes" ($k = 4$).

Focusing on responses to a single item in the questionnaire, the issue immediately arises of synthesizing the corresponding distribution of relative frequencies (shown in Table 1) by appropriate measures of location and dispersion.

Table 1. Distribution of relative frequencies.

Ordinal categories	m_1	…	m_i	…	m_k
Relative frequencies	p_1	…	p_i	…	p_k

The discussion of this issue, introduced in Section 2, is developed in Sections 3, while Section 4 extends the study to multidimensional measures.

2. Location and dispersion of ratings

In order to take into account the ordinal nature of student ratings, it is appropriate to represent the standard simplex of frequency distributions with *k* categories

$$\Delta_k = \left\{ \mathbf{p} = (p_1,\ldots,p_k) \in [0,1]^k : \sum_{i=1}^{k} p_i = 1 \right\} \tag{1}$$

as $\Delta'_k = \{(F'_1,F'_2,\ldots,F'_k) : 1 = F'_1 \geq F'_2 \geq \ldots \geq F'_k \geq 0\}$ using the relative inverse cumulative frequencies $F'_i = \sum_{j=i}^{k} p_j$ (i = 1,…,k). It is also appropriate to require that a location index

$$L : \Delta'_k \to \{m_1,\ldots,m_k\} \tag{2}$$

preserve the stochastic ordering, i.e. if

$$(F'_1,\ldots,F'_k),(G'_1,\ldots,G'_k) \in \Delta'_k$$

with

$$F'_1 \leq G'_1,\ldots,F'_k \leq G'_k,$$

then

$$L(F'_1,\ldots,F'_k) \leq L(G'_1,\ldots,G'_k).$$

Such an index, in fact, reaches its minimum or maximum value, respectively, in the case of complete dissatisfaction (p_1 = 1) or total satisfaction (p_k = 1):

$$L(F'_1,\ldots,F'_k) = \begin{cases} \min(L) : F'_2 = 0 \\ \max(L) : F'_k = 1 \end{cases}$$

In addition, such an index of location, as the median $me = \min\{m_i : F'_{i+1} \leq 0.5\}$, can be associated with a suitable measure of ordinal dispersion, as the index of Leti

$$G_{ord} = \sum_{i=1}^{k} [F_i(1-F_i) + F'_i(1-F'_i)] \in [0,(k-1)/2] \tag{3}$$

or the ordinal entropy

$$H_{ord} = -\sum_{i=1}^{k}\left[F_i \log(F_i) + F_i' \log(F_i')\right] \in \left[0, (k-1)\log(2)\right] \qquad (4)$$

These indexes allow to decompose the overall dispersion measured in a set of C University courses in two parts, namely "between" and "within" the courses:

$$G_{ord} = 2\sum_{i=1}^{k-1} F_i(1-F_i) = G_{within} + G_{between}$$

(Grilli and Rampichini, 2002) and

$$H_{ord} = -\sum_{i=1}^{k-1}\left[F_i \log(F_i) + (1-F_i)\log(1-F_i)\right] = H_{within} + H_{between},$$

where

$$G_{within} = 2\sum_{c=1}^{C}\pi_c \sum_{i=1}^{k-1} F_{i|c}(1-F_{i|c})$$

$$G_{between} = 2\sum_{c=1}^{C}\pi_c \sum_{i=1}^{k-1}(F_{i|c} - F_i)^2$$

$$H_{within} = -\sum_{c=1}^{C}\pi_c \sum_{i=1}^{k-1}\left[F_{i|c}\log(F_{i|c}) + (1-F_{i|c})\log(1-F_{i|c})\right]$$

$$H_{between} = -\sum_{c=1}^{C}\pi_c \sum_{i=1}^{k-1}\left[F_{i|c}\log\frac{F_{i|c}}{F_i} + (1-F_{i|c})\log\frac{1-F_{i|c}}{1-F_i}\right]$$

and $F_i = \sum_{h=1}^{i} p_h = \sum_{c=1}^{C}\pi_c F_{i|c}$ (*i* = 1,…,*k*) is the relative cumulative frequency expressed as a finite mixture of relative cumulative frequencies $F_{i|c}$, corresponding to ratings of the course *c*, with weights π_c given by the proportions of students who evaluate the course *c* (*c* = 1,…,C). In this regard, we note that the higher is the value of the "between" component, the more is the influence of the course features on the student ratings.

In general, an index of ordinal dispersion *D* attains its minimum if

$$\exists i \le k \quad p_i = 1$$

and attains its maximum in the case of polarization

$$p_1 = \tfrac{1}{2} = p_k$$

(Leti, 1983). This happens for the indexes (3) and (4).

In particular, it follows that $H_{ord} \ge G_{ord}$ and the index (4), when compared with (3), gives more weight to the frequencies

$$F_1 = p_1, F_k' = p_k$$

when they are small.

In the case of four courses and $k = 4$, Table 2 illustrates how the above approach produces results easy to communicate, despite being characterized by a lack of ability to discriminate situations very different from each other, because of the poor range in (2).

Table 2. Some traditional measures of location and ordinal dispersion in the case of $k = 4$.

	m_1	m_2	m_3	m_4	Location		Dispersion	
Course	p_1	p_2	p_3	p_4	*Mode*	*Median*	*Leti*	*Entropy*
1	0.10	0.20	0.30	0.40	m_4	m_3	1.08	1.61
2	0.01	0.20	0.39	0.40	m_4	m_3	0.83	1.24
3	0.01	0.10	0.49	0.40	m_3	m_3	0.70	1.08
4	0.10	0.30	0.30	0.30	nd	m_3	1.08	1.61

Table 2 also shows that the mode, as location index, does not preserve the stochastic ordering.

3. Satisfaction indexes

From a general point of view, a satisfaction index can be identified with a real-valued function defined on the standard simplex (1)

$$S : \Delta_k \to \Re \tag{5}$$

for which the following two conditions apply.

Condition of strict monotony

The function S is said to be strictly monotone if it is strictly increasing with respect to the stochastic ordering, that is to say

$$\forall j \le k \;\; \forall i < j \;\; \forall p \in]0, p_i] \quad S(p_1,\ldots,p_i,\ldots,p_j,\ldots,p_k) < S(p_1,\ldots,p_i - p,\ldots,p_j + p,\ldots,p_k) \tag{6}$$

Being a reinforcement of stochastic order preserving condition, (6) implies the following condition

$$S(p_1,\ldots,p_k) = \begin{cases} \min(S) : p_1 = 1 \\ \max(S) : p_k = 1 \end{cases} \tag{7}$$

Condition of equivariance

The function (5) is said to be equivariant if it is equivariant with respect to the order reversing permutation of $(p_1, p_2,\ldots,p_{k-1}, p_k)$, that is

$$S(p_1, p_2,\ldots,p_{k-1}, p_k) - \min(S) = \max(S) - S(p_k, p_{k-1},\ldots,p_2, p_1). \tag{8}$$

This condition identifies the value *s* of (5) corresponding to any frequency distribution invariant under order reversing permutation, i.e. such that

$$(p_1, p_2, \ldots, p_{k-1}, p_k) = (p_k, p_{k-1}, \ldots, p_2, p_1),$$

since from (8) follows immediately $s = [\min(S) + \max(S)]/2$, indicating a medium level of satisfaction.

The following are some examples of satisfaction indexes:

$$CG = 2p_1 + 5p_2 + 7p_3 + 10p_4$$

(k = 4), due to Chiandotto et al. (2000) and used in teaching evaluation by Universities of Cagliari, Ferrara, Florence, Macerata, Milan, Palermo, Sannio Sassari, Trieste and Urbino;

$$M_u = -p_1 - up_2 + up_3 + p_4$$

(k = 4, 0 < u < 1), studied by Marasini et al. (2011) and used by Universities of Pavia and Insubria;

$$SDI = \sum_{i=1}^{k} F_i' = \sum_{i=1}^{k} ip_i$$

proposed by Cerchiello et al. (2010) and employed by Universities of Bergamo, Brescia, Naples East, Molise, Rome - Tor Vergata, Turin, Venice - Ca 'Foscari and Verona.

However, the index due to Capursi et al. (2001)

$$IS_r = 1 - \left(\frac{1}{k-1} \sum_{i=1}^{k-1} F_i^r \right)^{\frac{1}{r}} \in [0,1]$$

(r > 0) is strictly monotone, but is not equivariant, except in the case r = 1; while the index introduced by Civardi et al. (2006) at University of Milano – Bicocca

$$CI_v = \frac{p_4 + p_3 - p_2 - p_1 + v\left(\frac{p_4}{p_4 + p_3} - \frac{p_1}{p_2 + p_1} \right)}{1 + v} \in [-1,1]$$

(k = 4, 0 < v ≤ 1) meets the condition of equivariance and (7), but strict monotony is not met.

In the case of four courses and k = 4, Table 3 illustrates the above considerations. For instance, the first course and the third course show that CI_v, with v = 1, is not strictly monotone, whereas the first course and the last course show that the index

$$IS_{0.5} = 1 - \left(\frac{1}{3} \sum_{i=1}^{3} \sqrt{F_i} \right)^2 \in [0,1]$$

is not equivariant (if this index were equivariant, then we would have 0.30 instead of 0.35 in the last row of the Table 3).

Table 3. Some popular measures of student satisfaction in the case of $k = 4$.

Course	p_1	p_2	p_3	p_4	CG	$M_{0.5}$	SDI	$IS_{0.5}$	CI_1
1	0.10	0.20	0.30	0.40	7.30	0.35	3.00	0.70	0.32
2	0.00	0.30	0.30	0.40	7.60	0.40	3.10	0.81	0.49
3	0.10	0.00	0.50	0.40	7.70	0.55	3.20	0.78	0.12
4	0.40	0.30	0.20	0.10	4.70	-0.35	2.00	0.35	-0.32

In general, the weighted mean

$$S_{\mathbf{w}} = \sum_{i=1}^{k} w_i p_i \quad (9)$$

is a function defined on the standard simplex, which is strictly monotone, if it satisfies the inequalities

$$w_1 < w_2 < \ldots < w_{k-1} < w_k , \quad (10)$$

and equivariant, if it also satisfies the equalities

$$w_i - w_1 = w_k - w_{k+1-i} \quad (11)$$

($i = 1,\ldots,k$). Thus the linear function (9) with weights satisfying conditions (10) and (11) can be called linear satisfaction index (being strictly monotone and equivariant).
More specifically, each linear satisfaction index is completely determined by a weight vector $\mathbf{w} = (w_1,\ldots,w_k)$ satisfying (10) and (11), and sends the generic element $\mathbf{p} = (p_1,\ldots,p_k)$ of the standard simplex (1) to the point defined by the inner product $\mathbf{p} \cdot \mathbf{w} = \sum_{i=1}^{k} p_i w_i \in [w_1, w_k]$.
Such a point, determined on the basis of the k coordinates $p_1,\ldots,p_k$, can also be represented as

$$S_{\mathbf{w}} = (1 - S_{\tilde{\mathbf{w}}}) w_1 + S_{\tilde{\mathbf{w}}} w_k = w_1 + S_{\tilde{\mathbf{w}}} (w_k - w_1)$$

by means of the coordinate

$$S_{\tilde{\mathbf{w}}} = \frac{S_{\mathbf{w}} - w_1}{w_k - w_1} \in [0,1]$$

defined through standardization. In this way, to every satisfaction index $S_{\mathbf{w}}$ there correspond a standardized index

$$S_{\tilde{\mathbf{w}}} = \sum_{i=1}^{k} \tilde{w}_i p_i ,$$

where the weights

$$\tilde{w}_i = \frac{w_i - w_1}{w_k - w_1} = \frac{w_k - w_{k+1-i}}{w_k - w_1} \in [0,1]$$

(i = 1,…,k) satisfy (10) and (11). Moreover, we can see that (11) takes the simple form $\widetilde{w}_{k+1-i} = 1 - \widetilde{w}_i$.

It is interesting to note that some linear satisfaction indexes can be interpreted in terms of probability. To see this, let X and Y be random variables providing ratings of a student chosen at random, respectively, from the observed distribution $p_1,\ldots,p_k$ and from a virtual distribution we require to be invariant under order reversing permutation. It is thus possible to define the probability that a student randomly selected from the observed distribution provides a rating not less than the virtual rating drawn from an invariant distribution corresponding to a medium level of satisfaction:

$$P(X \geq Y) = \sum_{i=1}^{k} P(Y \leq m_i) p_i \ .$$

Similarly, we may define the difference $P(X > Y) - P(X < Y)$.

For instance, if we choose $\left(\frac{1}{k},,\ldots,\frac{1}{k}\right)$ as the invariant distribution of Y, we get

$$P(X \geq Y) = \sum_{i=1}^{k} \frac{i}{k} p_i = \frac{SDI}{k}$$

and

$$P(X > Y) - P(X < Y) = \sum_{i=1}^{k} \frac{2i-k-1}{k} p_i = \frac{2SDI-k-1}{k},$$

which for k = 4 becomes

$$P(X > Y) - P(X < Y) = -\tfrac{3}{4} p_1 - \tfrac{1}{4} p_2 + \tfrac{1}{4} p_3 + \tfrac{3}{4} p_4 = \tfrac{4}{3} M_{1/3}.$$

However, if Y has invariant distribution $\left(0,\frac{1}{2},\frac{1}{2},0\right)$, then

$$P(X \geq Y) = \tfrac{1}{2} p_2 + p_3 + p_4$$

is a weighted mean that does not meet the conditions of strict monotony and equivariance, while

$$P(X > Y) - P(X < Y) = -p_1 - \tfrac{1}{2} p_2 + \tfrac{1}{2} p_3 + p_4 = M_{1/2}.$$

A further advantage of the linear satisfaction indexes is that we can measure the ordinal dispersion (in the sense of Section 2) by means of the variance

$$D_{\mathbf{w}} = \sum_{i=1}^{k} w_i^2 p_i - \left(\sum_{i=1}^{k} w_i p_i\right)^2 \in \left[0, \left(\frac{w_k - w_1}{2}\right)^2\right],$$

which can be decomposed as

$$D_{\mathbf{w}} = \sum_{i=1}^{k} (w_i - S_{\mathbf{w}})^2 p_i = D_{within} + D_{between},$$

where

$$D_{within} = \sum_{c=1}^{C} \pi_c \sum_{i=1}^{k} \left(w_i - S_{\mathbf{w}|c}\right)^2 p_{i|c}$$
$$D_{between} = \sum_{c=1}^{C} \pi_c \sum_{i=1}^{k} \left(S_{\mathbf{w}|c} - S_{\mathbf{w}}\right)^2$$

and

$$p_i = \sum_{c=1}^{C} \pi_c p_{i|c}$$

(i = 1,…,k) is the relative frequency expressed as a finite mixture of relative frequencies $p_{i|c}$, corresponding to ratings of the course c, with weights π_c given by the proportions of students who evaluate the course c, and $S_{\mathbf{w}|c}$ is the linear satisfaction index for the course c (c = 1,…,C).

4. Multidimensional satisfaction indexes

If we extend our view to include C courses evaluated by a questionnaire with Q questions, we may construct a multidimensional index of the form

$$MS = \sum_{c=1}^{C} \pi_c MS_c$$

(Russo, 2002), being π_c the proportions of students who evaluate the course c (c = 1,…,C) and

$$MS_c = \frac{\sum_{q=1}^{Q} \left(1 - \widetilde{D}_{cq}\right) \widetilde{S}_{cq}}{\sum_{q=1}^{Q} \left(1 - \widetilde{D}_{cq}\right)},$$

where, referring to the question q (q = 1,…,Q) about the course c, $\widetilde{S}_{cq}$ and $\widetilde{D}_{cq}$ represent, respectively, a standardized satisfaction index and a standardized ordinal dispersion index (in the sense of Section 2).

5. Conclusions

The paper shows how to measure student satisfaction through real-valued functions satisfying some appropriate conditions. In this context, further work will involve providing a complete characterization of satisfaction measures and a general representation of them.

References

1. Capursi, V. and Porcu, M. **La didattica universitaria valutata dagli studenti: un indicatore basato su misure di distanza fra distribuzioni di giudizi**. In: Atti Convegno Intermedio della Società Italiana di Statistica "Processi e Metodi Statistici di Valutazione", Roma 4-6 giugno 2001, 2001
2. Cerchiello, P., Dequarti, E., Giudici, P. and Magni C. **Scorecard models to evaluate perceived quality of academic teaching**. Statistica & Applicazioni, 2, pp. 2010, pp. 145-155
3. Chiandotto, B. and Gola, M. **Questionario di base da utilizzare per l'attuazione di un programma per la valutazione della didattica da parte degli studenti**. Rapporto di Ricerca RdR1-00. Gruppo di Ricerca MIUR-CNVSU, 2000
4. Civardi, M., Crocetta, C. and Zavarrone E. **Summary indicators of opinion expressed by the users of a given service**. Statistica. LXVI (4), 2006, pp. 373-388
5. Giovagnoli, A. and Regoli, G. **Representation Theorems for Measures of Location and for Measures of Dispersion**, Stochastic Orders and Decision under Risk. Lecture Notes-Monograph Series, Vol. 19, 1991, pp. 141-156
6. Grilli, L. and Rampichini, C. **Scomposizione della dispersione per variabili statistiche ordinali**. Statistica. LXII (1), 2002, pp. 111-116
7. Leti, G. **Statistica descrittiva,** il Mulino, Bologna, 1983
8. Russo, M.A. **Indicatori statistici per la valutazione della qualità della didattica universitaria: una proposta metodologica**. Statistica. LXII (3), 2002

---

[1] Since 1998 Full Professor in Statistics at University of Milan – Bicocca, Department of Statistics. From 1990 to 1997 Full Professor in Statistics at University of Milan. From 1986 to 1989 Full Professor in Statistics at University of Brescia.

[2] Since 2007 Associate Professor in Statistics at University of Milan – Bicocca, Department of Statistics. From 2002 to 2007 Researcher in Statistics at University of Milan – Bicocca, Department of Statistics. In 2002 PhD in Statistics at University of Milan - Bicocca. From 2001 to 2002 Teacher of Mathematics. In 2001 Qualification in Mathematics.