

On the Choice between the Exponential and the Normal in the Indirect Scaling

*Sulla Scelta tra l'Esponenziale e la Normale
nella Quantificazione Determinata Indiretta*

Giovanni Portoso

Facoltà di Economia - via Perrone, 18 - Novara - portoso@eco.unipmn.it

Riassunto : Oltre alla normale standardizzata, di solito assunta come *variabile latente* nella *quantificazione determinata indiretta*, vengono esaminate l'esponenziale negativa e positiva standardizzate insieme ad un apposito indicatore, normalizzato tra -1 e $+1$, che è d'aiuto nella scelta. Un'applicazione condotta sui giudizi espressi dagli studenti nei riguardi di alcuni servizi erogati dalla Facoltà di Economia di Novara, consente di validare sia la scelta dell'esponenziale negativa o positiva in alternativa alla normale sia il funzionamento dell'indicatore.

Keywords : Normal and exponential distribution, indirect scaling.

1. Introduction

The indirect scaling, supposing that a continuous *latent variable* is included in the empirical distribution, tends to convert on metric scale the ranked categories, through what the judgments of validity on the products or the services are expressed. After fixing the equivalence between the distribution function of a *latent variable* adopted and the cumulative function observed relative to the categorical ordinal variable containing the judgments expressed, the used methodology considers the progressive areas accumulated of the distribution function observed and it deduces the theoretical quantiles associates to the chosen function (Delvecchio, 1996; De Luca, 2002). Particularly, calling with $F(i_2)$ the theoretical cumulative distribution function assumed and putting the equivalence with the empirical one, $F(i)$, we have :

$$F(i_2) = F(i-1) + f(i)/2 \quad (1)$$

where $f(i)$ is the frequency observed of the modality i ; if the normal standardized is assumed as *latent variable*, the quantile, z_i , is obtained as it follows:

$$z_i = \Phi^{-1}[F(i_2)] \quad (2)$$

The procedure to calculate the quantiles combined to the modalities, doesn't change, essentially, if another latent variable is considered. The generalized use of the normal standardized, not always, results able to justify the process of quantitative conversion, above all, when there are empirical frequencies sidely thickened at the extreme points. In these cases, the assumption of the normal can lead to absurd standardized values, that, sometimes, overturn the results that can become negative from positive and viceversa (Portoso, 2003a). We proposed the alternative use of the standardized

negative exponential, above all, when the frequencies figure thickened on the first modalities and the use of the exponential positive when the frequencies are condensed mostly on the right tail of the distribution. Therefore we considered a new indicator, opportunely normalized, that can be useful when there are situations of perplexity in the choice between the above mentioned variables. It has been carried out an interesting application of this methodology to the valuation of the university services of the Novara Faculty of Economics.

2. The negative and positive exponential standardised

The negative exponential , X , defined as density function, as :

$$f(x) = \exp(-x) \quad \text{for } (0 \leq x \leq +\infty) ; \quad f(x) = 0 \quad \text{elsewhere;}$$

with unitary mean and variance (Portoso, 2003a), has been standardized in S as it follows:

$$S = X - 1 \quad \text{with} \\ f(s) = \exp(-s-1) \quad \text{for } (-1 \leq s \leq +\infty) ; \quad f(s) = 0 \quad \text{elsewhere;}$$

and distribution function, $F(s)$, as :

$$F(s) = \int_{-1}^s \exp(-s-1) ds = 1 - \exp(-s-1) \quad (3)$$

Referring to the equality, of which (1) , centered on the modality i , we have :

$$F(i_2) = 1 - \exp(-s_i-1) \quad \text{therefore : } s_i = -1 - \ln[1 - F(i_2)] \quad (4)$$

where s_i indicates the quantile of the distribution function of the negative standardized exponential, associated to the modality i .

The positive exponential, Y , defined as density function, as :

$$f(y) = \exp(y) \quad \text{for } (-\infty \leq y \leq 0) ; \quad f(y) = 0 \quad \text{elsewhere;}$$

with unitary mean and variance, has been standardized in P as it follows:

$$P = [Y - (-1)] = Y + 1 \quad \text{with} \\ f(p) = \exp(p-1) \quad \text{for } (-\infty \leq p \leq 1) ; \quad f(p) = 0 \quad \text{elsewhere;}$$

and distribution function, $F(p)$, as :

$$F(p) = \int_{-\infty}^p \exp(p-1) dp = \exp(p-1) \quad (5)$$

Referring to the equality, of which (1) , centered on the modality i , we have :

$$F(i_2) = \exp(p_i-1) \quad \text{therefore : } p_i = 1 + \ln[F(i_2)] \quad (6)$$

where p_i indicates the quantile of the distribution function of the positive standardized exponential, associated to the modality i .

3. The choice index between exponential and normal

We suppose that an indicator, quantifying unilateral accumulations on the left or the right tail of a frequencies distribution, associated to categorical modalities of rectilinear type, is revealed somewhat opportune if it is standardized between -1 and $+1$ and if it assumes value 0 when a balance condition of the codified values (with constant step) of the modalities, based on the center of the distribution - but not necessarily connected with a symmetry of the frequencies - exists. An indicator, EN , that answers to these conditions, is :

$$EN = \sum_{i=1}^{[s/2]} (f_{s-i+1} - f_i) * (s-2i+1) / (s-1) \quad (7)$$

where $[s/2]$ is to indicate the maximum integer contained in $s/2$, f_i is the relative frequency associated to the modality i , f_{s-i+1} is combined with the opposed modality and $s-2i+1$ is the gradual distance, that is the difference of position between the modality i and the speculate $[(s-i+1)-i]$; $(s-1)$ is the maximum value that the numerator can assume (Portoso, 2003b).

4. The empirical analysis

In the Table 1, for their representative effects, we choiced some frequencies distributions concerning the judgments expressed by students, at may 2003, to the services supplied by the Novara Faculty of Economics and considered in a degree thesis (Castellaneta A.).

Table 1 : *Frequencies distributions of judgments expressed by students towards to the some university services of the Novara Faculty of Economics with the corresponding index EN, quantiles and linear transformations*

JUDGMENTS		SERVICES								
		I	II	III	IV	V	VI	VII	VIII	IX
VERY INSUFFIC.		9,3	17,5	38,9	4,2	3,2	7,1	1,6	0,9	0,5
INSUFFIC.		21,8	21,8	27,3	14,1	9,8	12,5	4,7	1,0	6,8
MEDIocre		26,4	22,0	16,4	23,2	22,8	12,7	14,0	11,4	14,8
SUFFIC.		22,5	18,6	11,2	28,5	27,5	20,8	22,3	19,3	29,8
QUITE SUFFIC.		13,2	10,8	4,6	18,2	20,8	19,7	24,4	22,6	22,8
GOOD		6,6	8,7	1,6	10,9	14,8	22,6	27,7	34,6	22,0
EXCELLENT		0,2	0,6	0,0	0,9	1,1	4,6	5,3	10,2	3,3
TOTALS		100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
INDICATOR EN		-0,24	-0,29	-0,52	-0,07	0,01	0,07	0,23	0,35	0,16
QUANT. = QM	NORM.	0,006	0,014	0,042	-0,000	-0,004	-0,005	-0,011	-0,013	-0,006
	EXP.	-0,030	-0,026	-0,023	-0,034	0,030	0,032	0,023	0,028	0,031
TRANSF. = TL	NORM.	6,36	6,84	8,53	5,98	5,76	5,73	5,34	5,23	5,67
	EXP.	4,21	4,43	4,62	3,99	7,80	7,89	7,40	7,69	7,84

We selected the distributions in order to get a band fairly wide of the cases; in fact the first three distributions contain frequencies crowded especially on the categories

considered negative; the next three presents the frequencies concentrated on the central categories whereas the last three are on the others considered positive.

The indicator *EN* seems to suggest, for the first three distributions, the adoption of the negative exponential; for the others three (IV, V, VI) the normal and, for the last three, the positive exponential. Instead of considering the values of the medium quantiles, *QM*, obtained pondering the quantiles of category with the respective frequencies, it seems more useful to pay attention to the values of the linear transformed ones *TL*, for $TL=6+60*QM$, whose examination allows to find the incongruities connected with the indiscriminate assumption of the standardized normal. In fact the use of the normal, for the first three distributions, would lead to one positive quantification whereas, considering last the three, it would arrive to a negative quantification, both paradoxical; the adoption of the exponentials leads, instead, to the results that fit better the distributive structures of reference; viceversa, for the distributions IV, V and VI, the normal appears perfect.

It is obvious that, by the calculation of further means, it is possible to value macro-services if the ones, that are examined, can be included in the last cases .

In order to facilitate the observation, we indicated the choices, considered the best, with the bond font whereas the italics one marks the values reputed unacceptable.

5. Conclusions

The analysis evidences that the alternative use of exponential - rather than the normal in the distributions with unilateral accumulations of frequencies - avoids the manifestation of absurd situations while the indicator *EN* carries out an effective action of support in the *latent variable* choice. In particular the values of *EN*, included in the range $-0,1$ and $+0,1$, direct to the normal while the values, outward of the above mentioned limits, are inclined to the negative or positive exponential according to the sign. It must be emphasize that the examined cases, in this work or in previous ones, leave ambivalent solutions for values near the aforesaid limits.

At last it is found out that the adoption of a linear transformation of the quantiles rather than one another, answers to discretionary criteria of the operator and however it is connected to the adopted monotone syntagmatic scale.

References

- Castellaneta A., *Indagine statistica sulla Customer Satisfaction degli studenti sui servizi della Facoltà di Economia di Novara*, Tesi di laurea in Statistica, RBelatore G. Portoso, Università degli studi del Piemonte Orientale “A. Avogadro”, a.a. 2002-03.
- De Luca A. (2002), Le applicazioni dei metodi statistici alle analisi di mercato, *Manuale di Marketing Quantitativo*, Franco Angeli, Milano.
- Delvecchio F. (1996), *Statistica per la ricerca sociale*, Cacucci, Bari.
- Portoso G. (2003a) La quantificazione determinata indiretta nella customer satisfaction : Un approccio basato sull’uso alternativo della Normale e dell’Esponenziale, *Quaderni del Dipartimento SEMEQ dell’Univ.à del Piemonte Orientale*, 53, pagg. 1-14.
- Portoso G. (2003b), Un indicatore di addensamento codale di frequenze per variabili categoriche ordinali basate su giudizi, *Quaderni del Dipartimento SEMEQ dell’Univ.à del Piemonte Orientale*, 66, pagg. 1-18.