

Un'applicazione MDS all'analisi di dati sensoriali

MDS application in sensorial data analysis

Eugenio Brentari

Università di Brescia, Dipartimento Metodi Quantitativi,
C.da S. Chiara 50, 25122 Brescia, e-mail: brentari@eco.unibs.it

Abstract: *Vinitaly*, one of the most important international fairs of the wine industry which takes place in Verona, offers the pleasure of a tasting desk with grappa and spirits. In this paper the sense faculties of the tasters to distinguish Italian grappa from some Brazilian *cachaça* are analysed through the multidimensional scaling.

Keywords: Multidimensional Scaling (MDS), Sensorial Analysis

Introduzione

Nell'analisi sensoriale (insieme di tecniche e metodi che permettono di misurare, attraverso gli organi di senso, quanto viene percepito di un prodotto o servizio) si è soliti distinguere tra *test di laboratorio* e *test sul consumatore*. Nel primo, in genere, si seleziona un numero ridotto di giudici con abilità sensoriali note, mentre il secondo solitamente si effettua durante manifestazioni fieristiche in cui è possibile far avvicinare al banco di assaggio un numero elevato di visitatori. Da anni, in occasione di *Vinitaly*, una delle più importanti manifestazioni internazionali del settore vitivinicolo che si svolge a Verona, viene effettuato un test sul consumatore rivolto alle grappe ed alle bevande "spiritose". Nell'edizione 2001 tra i vari prodotti all'assaggio erano presenti anche tre *cachaça* brasiliane, assimilabili alla grappa italiana ma derivate da materie prime diverse. Nel presente lavoro si analizza, attraverso l'impiego dello *scaling* multidimensionale, come gli assaggiatori hanno reagito alla presenza di questo diverso prodotto.

1. Lo scaling multidimensionale: cenni metodologici

Lo *scaling multidimensionale* (MDS) è l'insieme delle procedure per l'analisi di una o più matrici di "indici di prossimità" rappresentanti la somiglianza o la dissomiglianza tra tutte le possibili coppie di n unità statistiche (Zani, 1999).

I dati di *input* possono essere raccolti tramite un *modello metrico* o un *modello non metrico*: nel primo caso le misure di prossimità sono espresse almeno su scala di intervalli; nel secondo caso si considera una semplice scala ordinale e la soluzione si ottiene attraverso un procedimento iterativo per approssimazioni successive.

A seconda del modo in cui vengono raccolti i dati si distingue tra misure di *prossimità diretta* e misure di *prossimità indiretta*. La procedura più comune per ottenere le prime è quella di chiedere ad una o più persone di esprimere direttamente un giudizio sulla diversità percepita tra oggetti. Le misure di prossimità indiretta sono invece calcolate in funzione delle modalità di caratteri quantitativi o qualitativi rilevati per ciascuna delle unità statistiche; operativamente si dispone di una matrice del tipo unità per variabili e l'*output* dell'analisi è costituito da una matrice \mathbf{X} , di dimensioni $n \times k$, che consente di effettuare una rappresentazione grafica dello *scaling multidimensionale*. Le unità, infatti, possono essere rappresentate in uno spazio a k dimensioni nel quale la generica

i -esima unità ha la posizione individuata dal vettore di coordinate $x_i = (x_{i1}, x_{i2}, \dots, x_{ik})$ che costituisce la i -esima riga della matrice \mathbf{X} .

Abitualmente si considera un numero ridotto di dimensioni, due o tre al massimo, verificando la bontà dell'adattamento di tale soluzione. A tale scopo, Kruskal (1964) ha proposto una misura, nota come *stress* (*Standardized Residual Sum of Squares*) che è interpretabile come una sorta di varianza residua, perché esprime quanta parte della variabilità delle prossimità osservate non è spiegata dalla configurazione. L'indice assume quindi valore minimo zero quando l'adattamento tra la soluzione ottenuta ed i dati originali è perfetto, mentre assume valore massimo uno, quando si riscontra una completa mancanza di adattamento. Sono considerati insufficienti valori dell'indice superiori al 20%, discreti valori compresi tra il 10% ed il 20%, buoni valori compresi tra il 5% ed il 10% ed eccellenti valori inferiori al 5%.

2. Grappa & C. Tasting

I dati utilizzati in questo lavoro sono stati raccolti nel corso della 35^a edizione di *Vinitaly*, svoltasi nell'aprile del 2001 e promossa dall'Ente Fiera di Verona. Da anni, nel corso di questa importante manifestazione internazionale viene effettuato un test sul consumatore, noto come "*Grappa & C. Tasting*", attraverso un banco di assaggio di grappe, acquaviti e liquori, sotto il patrocinio dell'Istituto Nazionale Grappa e curato dal CSFA di Brescia⁽¹⁾.

Nel caso in esame erano presenti circa 150 prodotti, tra cui 25 grappe invecchiate o affinate, 100 grappe giovani o giovani aromatiche e 3 *cachaça*, oltre a vari altri distillati e liquori. Proprio la presenza della *cachaça* (*caninha* o *pinga*), tipica bevanda brasiliana, ottenuta dalla distillazione della canna da zucchero, a differenza della grappa che è invece un distillato di vinaccia, rende interessante la possibilità di applicare la tecnica MDS per analizzare la reazione degli assaggiatori di fronte a questo distillato.

Hanno effettuato la degustazione 1122 persone (di cui l'88% maschi) per un totale di 4804 assaggi. A ogni partecipante è stata consegnata una scheda in cui, oltre ad alcune informazioni di carattere generale (sesso, classe d'età, provenienza, professione, titolo di studio, ...), sul grado di competenza nel giudizio e sulla tipologia di consumatore del prodotto, si è chiesto di esprimere un giudizio sulle bevande assaggiate, valutando i prodotti degustati sotto tre profili: *visivo*, *olfattivo* e *gusto-olfattivo* secondo la seguente scala (modello non metrico e misure di prossimità indiretta):

1 = pessimo 2 = scarso 3 = mediocre 4 = discreto 5 = buono 6 = ottimo.

Per ciascuna grappa, acquavite o liquore si è considerata la frequenza di assaggio (*attraenza*) dato che al consumatore era data la possibilità di scegliere i prodotti da degustare da un apposito catalogo o direttamente dalle confezioni in mostra.

Dal totale dei prodotti iniziali, dopo aver scartato i liquori, i distillati d'uva e di frutta, per una corretta interpretazione dei dati si sono prese in considerazione solo le grappe con una certa *attraenza* (frequenze superiore a 8). In una prima analisi si sono accantonate anche le grappe "coperte" (*blind*), già presenti sul catalogo ma riproposte per ulteriori valutazioni senza fornirne origine e marca così da lasciare all'assaggiatore la possibilità di esprimere un giudizio sul prodotto affidandosi solo alle proprie percezioni senza condizionamenti esterni.

⁽¹⁾ Si ringrazia il *Centro Studi e Formazione Assaggiatori* (CSFA) di Brescia, diretto da Luigi Odello, per aver messo a disposizione i dati e le necessarie informazioni di supporto.

3. La valutazione degli assaggiatori

Dalle singole valutazioni degli assaggiatori si è ricavata la matrice delle valutazioni medie di ciascun prodotto rispetto ai tre profili esaminati. Ciò ha permesso di costruire una *matrice di prossimità*, determinata mediante la *distanza euclidea* tra ciascuna coppia di prodotti:

$$d_{ij} = \sqrt{(a_i - a_j)^2 + (b_i - b_j)^2 + (c_i - c_j)^2}$$

dove a_i e a_j rappresentano le valutazioni visive rispettivamente del prodotto i -esimo e del prodotto j -esimo; b_i e b_j le valutazioni olfattive; c_i e c_j le valutazioni gusto-olfattive, così da tenere conto di tutti e tre gli aspetti osservati.

Scelta la configurazione a due dimensioni, la rappresentazione di *multidimensional scaling* ottenuta è riportata nella Figura 1. L'indice di *stress* pari a 0,0687 denota il buon risultato dell'analisi. I punteggi medi delle valutazioni sono:

	visivo	olfattivo	gusto-olfattivo
	4,646	4,263	4,384

La maggior parte dei prodotti è contenuta in una nube di punti assai ristretta. Le grappe che hanno ottenuto una migliore valutazione su tutti e tre gli aspetti osservati si dispongono in alto a sinistra della nube (162, *grappa di Barbaresco* con punteggi elevati per tutti e tre gli aspetti e 155, *grappa di Schiava* valutata discretamente sotto l'aspetto gusto-olfattivo ma molto favorevolmente per gli altri due); al contrario, quelle che hanno riportato valutazioni basse si collocano in basso a destra (83, *grappa di Lambrusco* particolarmente penalizzata sotto tutti gli aspetti).

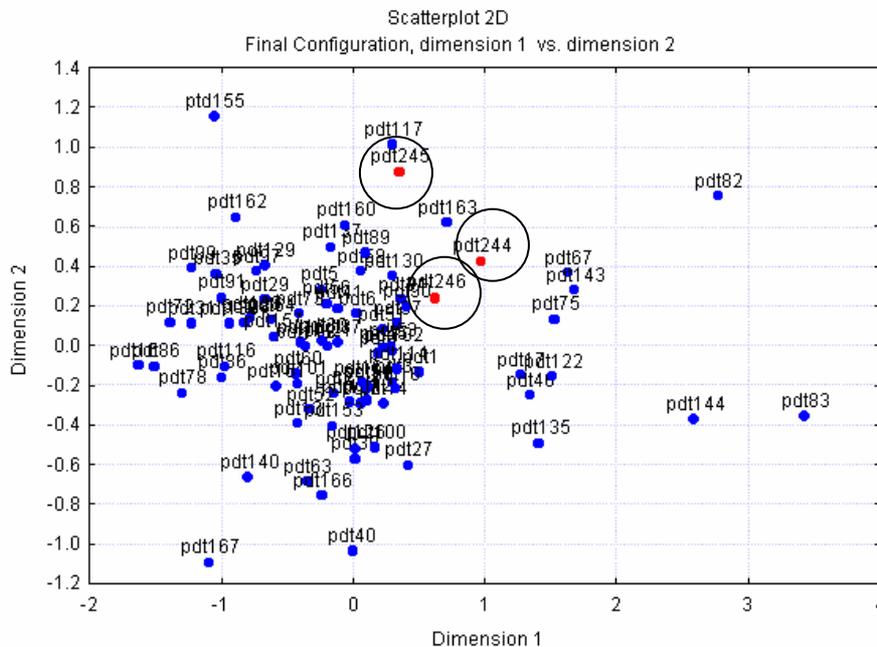


Figura 1: Configurazione in due dimensioni delle grappe esaminate

Le tre *cachaça* (244, 245, 246), proposte coperte, si collocano in una posizione, seppur di poco, superiore alla media delle grappe (nella figura sono cerchiare). Dall'analisi delle valutazioni medie ottenute da questa bevanda brasiliana

codice	visivo	olfattivo	gusto-olfattivo
244	4,415	4,122	3,805
245	4,618	4,539	3,868
246	4,396	4,229	4,043

risulta evidente che, mentre negli aspetti visivo e olfattivo le *cachaça* ottengono risultati più che soddisfacenti, in quello gusto-olfattivo hanno punteggi decisamente inferiori a quelli medi. Ciò ha consigliato di limitarne il confronto con le sole grappe invecchiate, che per colore e aroma più si avvicinano alla bevanda brasiliana. I risultati sono riportati nella Figura 2, dove è evidenziata la valutazione molto penalizzante delle *cachaça*. L'indice di *stress* denota, anche in questo caso, la bontà dell'adattamento (0,0523).

Non condizionati dall'aspetto visivo-olfattivo, i consumatori sono quindi in grado di riconoscere la diversità tra *cachaça* e grappa. Ciò sembrerebbe suggerire una ponderazione dei tre aspetti, privilegiando quello gusto-olfattivo che per questo prodotto è sicuramente quello più rilevante.

Altre interessanti analisi hanno riguardato il confronto grappe "coperte" e "scoperte", dove in alcuni casi è evidente l'effetto etichetta o l'influenza sulle valutazioni di alcune caratteristiche dei consumatori. Tenendo presente che non si trattava di esperti assaggiatori ma, al più, abituali consumatori di grappa, il giudizio sulle facoltà sensoriali di chi si è avvicinato al banco d'assaggio risulta nel complesso positivo.

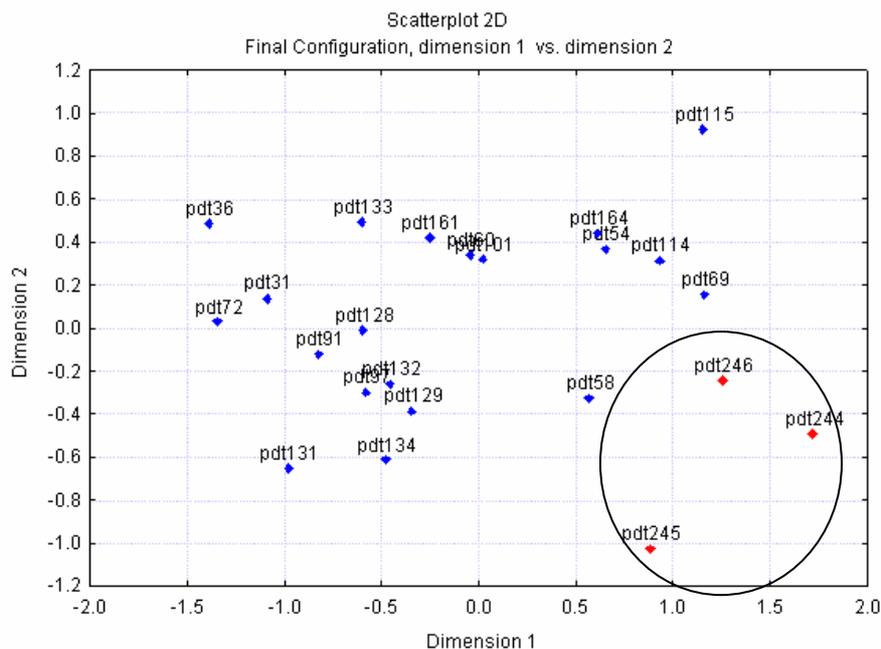


Figura 2: Confronto tra grappe invecchiate e *cachaça*

Bibliografia

- Coxon A.P.M. (1992), *The user's Guide to Multidimensional Scaling*, Heinemann Educational Books, London
- Everitt B.S. Rabe-Hesketh S.(1997), *The Analysis of Proximity Data*, Arnold, London
- Kruskal J.B. (1964), Multidimensional scaling by optimizing goodness-of-fit to a nonmetric hypothesis, *Psychometrika*, 29, 1-27
- Romano M.F. (1987), *Lo scaling multidimensionale: una bibliografia commentata con note introduttive*, Giunta Regionale Regione Toscana, Firenze
- Zani S. (1999), *Analisi dei dati statistici*, Vol. II, Giuffrè, Milano