La moyenne bipolaire pour l'analyse sensorielle

The bipolar mean in sensory analysis

Eugenio Brentari[1], Livia Dancelli[1], Walter Maffenini[2]

[1] *Dipartimento Metodi Quantitativi, Facoltà di Economia, Università degli Studi di Brescia*
E-mail: brentari@eco.unibs.it; dancelli@eco.unibs.it

[2] *Dipartimento di Metodi Quantitativi per le Scienze Economiche ed Aziendali, Facoltà di Economia, Università degli Studi di Milano-Bicocca*
E-mail: walter.maffenini@unimib.it

Résumé
Dans une publication récente, on a proposé comme mesure de synthèse pour les caractères qualitatifs ordinaux, la moyenne bipolaire qui est très utile dans le domaine de l'analyse sensorielle. Dans un travail suivant le concept de moyenne bipolaire a été appliqué aux caractères discrets et on a aussi proposé une mesure de variabilité associée à cette moyenne: l'écart moyen de la moyenne bipolaire. Cette mesure peut être calculée aussi pour les caractères qualitatifs ordinaux dont les catégories sont exprimées sur une échelle d'évaluation. L'écart moyen de la moyenne bipolaire est un indice absolu de variabilité; comme pour tous les indices de variabilité il est utile de définir sa valeur maximum dans le but de comparaison. On a donc déterminé la distribution de variabilité maximum et à partir de cette distribution, on a calculé la moyenne bipolaire correspondante. Dans cette étude, en ayant recours à différents types de données d'évaluation des produits alimentaires, on a approfondi la capacité de synthèse de la moyenne bipolaire en employant aussi la nouvelle mesure de dispersion.
Mots-clés: Moyenne bipolaire; dominance statistique; fréquences cumulées décroissantes; écart moyen de la moyenne bipolaire; maximum de l'écart moyen de la moyenne bipolaire; analyse sensorielle.

Abstract
In a recent work, the bipolar mean was proposed with the aim to summarize the ordinal variables. It is a synthetic distribution where the total size *n* is concentrated on one of the *k* categories of the variable or, at most, on two consecutive categories. The measure was derived according to the usual statistical dominance principle that is based on the retro-cumulative frequencies. Later the bipolar mean was extended to the discrete quantitative variables and a new variability measure was introduced, i.e. the mean deviation about the bipolar mean. This measure can be applied also to ordinal variables whose categories are expressed as scores on a numerical scale. Hence, the new way to summarize these variables can be useful in sensory analysis where it is often necessary to compare frequency distributions that represent the evaluations of judges or tasters in relation to some characteristics of different products. The assessment is usually based on simple synthetic measures such as the arithmetic mean or the median, but these indexes can provide contradictory answers.
The normalization of the mean deviation about the bipolar mean is proposed in this work. Moreover, some empirical evidences in sensory analysis are given with the

purpose of showing how the bipolar mean and the mean deviation about it can sometimes overcome the comparison problems.
Keywords: Bipolar mean; statistical dominance; retro-cumulative frequencies; mean deviation about the bipolar mean; maximum value of the mean deviation about the bipolar mean; sensory analysis.

1. Introduction

Dans l'analyse sensorielle il faut souvent confronter les distributions de fréquence des évaluations exprimées par n sujets (juges, dégustateurs, consommateurs) pour certains produits, sur une échelle de notation et sur certains caractéristiques (descripteurs).
On suppose (E. Brentari, L. Dancelli, 2008) que sept juges ont exprimé leurs évaluations, sur une échelle de notation de 1 (*très mauvais*) à 5 (*très bon*), par rapport à deux descripteurs (A et B) des trois "*grappe*" (G_1, G_2, G_3). Pour chaque descripteur on veut établir un classement des trois *grappe*. Normalement la comparaison est faite en utilisant de simples indices de synthèse tels que la médiane (M_e) ou la moyenne arithmétique (μ); la préférence va souvent à la médiane qui est moins sensible aux éventuelles notes anomales, c'est-à-dire les cas sporadiques qui sont évidemment discordants avec les évaluations exprimées par la majorité des juges.
Le classement créé par les deux indices n'est pas forcément le même (Tableau 1): pour M_e les trois *grappe* seraient jugées équivalentes dans le cas du descripteur A, par contre G_{3B} serait à préférer dans le cas du descripteur B; l'indice μ porterait à des conclusions opposées: les trois *grappe* seraient jugées équivalentes dans le cas du descripteur B, par contre G_{3A} serait à préférer dans le cas du descripteur A.

Notes	***Descripteur A***			***Descripteur B***		
	G_{1A}	G_{2A}	G_{3A}	G_{1B}	G_{2B}	G_{3B}
1	3	2	2	1	0	1
2	3	3	4	3	1	2
3	0	2	0	0	5	0
4	1	0	0	1	1	4
5	0	0	1	2	0	0
M_e	**2**	**2**	**2**	**2**	**3**	**4**
μ	**1,857**	**2**	**2,143**	**3**	**3**	**3**

Tableau 1: Distributions des évaluations de sept juges pour trois *grappe* sur deux descripteurs et indices correspondants

L'utilisation d'un nouvel indice de synthèse, la moyenne bipolaire, et de ses propriétés fournit la possibilité de sortir de l'impasse dans laquelle on peut se trouver dans des cas similaires à celui qu'on a présenté.
Dans une publication récente (W. Maffenini, M. Zenga, 2005) on a proposé, comme mesure de synthèse pour les caractères qualitatifs ordinaux, la moyenne bipolaire qui peut être utile (E. Brentari, L. Dancelli, 2005 et 2008) dans l'analyse sensorielle. Dans un travail de 2006 (W. Maffenini, Ma. Zenga) le concept de moyenne bipolaire a été appliqué aux caractères discrets et on a aussi proposé une mesure de variabilité associée à cette moyenne: l'écart moyen de la moyenne bipolaire. Cette mesure peut être calculée même pour les caractères qualitatifs ordinaux dont les catégories sont exprimées sur échelle d'évaluation en notes. L'écart moyen de la moyenne bipolaire est un indice absolu de variabilité; comme pour tous les indices de variabilité il est utile de définir sa valeur maximum dans le but de confrontation. On a donc défini (E. Brentari, L. Dancelli, W. Maffenini,

2009) la distribution de variabilité maximum et, après avoir déterminée la moyenne bipolaire correspondante, on a dérivée la valeur maximum de l'indice.
On va, maintenant, présenter brièvement les principaux aspects méthodologiques développés dans les travaux cités auparavant et auxquels on renvoie pour d'éventuels approfondissements.

2. La moyenne bipolaire

La moyenne bipolaire est une distribution qui concentre tous les n cas sur une seule modalité ou sur deux modalités contiguës et elle est cohérente avec le concept de dominance statistique qui s'appuie sur les fréquences cumulées décroissantes (ordre W). A chaque distribution empirique peut être associée la moyenne bipolaire correspondante qui est prise comme synthèse de la distribution vu qu'elle satisfait l'exigence d'ordre et possède diverses potentialités interprétatives.
Soit X une variable discrète qui prend les valeurs 1, 2, …, s, …, k et soient $n_1, n_2, \ldots, n_s, \ldots, n_k$ ($n = \sum_{s=1}^{k} n_s$) les fréquences associées aux k valeurs.
Soit $\mathcal{B}$ l'ensemble des toutes les distributions possibles qui satisfont les contraintes:

$$n_s \geq 0, \qquad s = 1, 2, \ldots, k; \qquad \sum_{s=1}^{k} n_s = n$$

dont le nombre est $\binom{n+k-1}{n}$.
Les distributions de $\mathcal{B}$ peuvent être comparées selon le principe de dominance statistique qui s'appuie sur les fréquences cumulées décroissantes:

$$R_S = \sum_{j=s}^{k} n_j \qquad s = 1, 2, \ldots, k$$

Si on a deux distributions $(n_1, n_2, \ldots, n_k)$ et $(n'_1, n'_2, \ldots, n'_k)$ la deuxième est dominée par la première selon l'ordre W, c'est-à-dire:

$$(n'_1, n'_2, \ldots, n'_k) \prec_W (n_1, n_2, \ldots, n_k)$$

si

$$R'_S \leq R_S \qquad s = 1, 2, \ldots, k$$

avec au moins une inégalité stricte.
En général il n'est pas possible ordonner selon W toutes les distributions appartenant à $\mathcal{B}$. Mais on peut concentrer l'attention sur son sous ensemble $\mathcal{B}^*$ qui est composé par les distributions pour lesquelles:

- *(i)* n se concentre sur une seule des k modalités de X;
- *(ii)* n se concentre sur deux modalités contiguës de X.

Le nombre des distributions qui composent $\mathcal{B}^*$ est $nk–n+1$.
Il est possible ordonner toutes les distributions de $\mathcal{B}$, et donc aussi les distributions qui n'appartiennent pas à $\mathcal{B}^*$, selon une fonction H, cohérente avec l'ordre W, tel que:

$$(n'_1, \ldots, n'_k) \prec_W (n_1, \ldots, n_k) \;\Rightarrow\; H(n'_1, \ldots, n'_k) < H(n_1, \ldots, n_k) \qquad (1)$$

La somme des fréquences cumulées décroissantes:

$$G = G\left(n_1,\ldots,n_k\right) = \sum_{s=1}^{k} R_s = \sum_{s=1}^{k}\sum_{j=s}^{k} n_j$$

est une fonction qui satisfait la (1).

Il est facile de vérifier que, en correspondance des *nk*–*n*+1 distributions appartenant à $\mathcal{B}^*$, ordonnées en sens croissant, la fonction *G* prend les valeurs suivantes en progression arithmétique de raison 1: *n*, *n*+1, …, 2*n*–1, 2*n*, 2*n*+1,…, *kn*–1, *kn*.

La valeur *n* est prise dans le cas de la distribution $(n, 0, \ldots, 0)$ et la valeur *kn* dans le cas de la distribution $(0, \ldots, 0, n)$. En outre les valeurs *n*, 2*n*, 3*n*,…, *kn* correspondent aux distributions qui concentrent *n* sur une seule modalité; par contre les valeurs qui restent correspondent aux distributions qui concentrent *n* sur deux modalités contiguës.

Il est possible de démontrer que en correspondance de toutes les distributions appartenant à $\mathcal{B}$ la fonction *G*:

(*i*) peut assumer seulement les valeurs entières dans l'intervalle [*n*, *kn*], c'est-à-dire les valeurs qui correspondent aux distributions appartenant à $\mathcal{B}^*$;

(*ii*) peut être exprimée comme:

$$G = \sum_{s=1}^{k} R_s = \sum_{s=1}^{k} s\, n_s \tag{2}$$

De la (*i*) on déduit que la fonction *G* attribue aux distributions de $\mathcal{B}$ qui n'appartiennent pas à $\mathcal{B}^*$ des valeurs qui sont dans le même intervalle [*n*, *kn*]. Il s'ensuit que l'ensemble $\mathcal{B}$ est partagé en *nk*–*n*+1 sous-ensembles (c'est-à-dire, le même nombre des distributions de $\mathcal{B}^*$). Aux distributions qui appartiennent au même sous-ensemble, qui comprend un seul élément de $\mathcal{B}^*$, la fonction *G* assigne la même valeur *g* et, donc, ces distributions sont équivalentes selon l'ordre *W*. Il est donc possible de représenter toutes les distributions du même sous-ensemble avec la seule distribution qui appartienne à $\mathcal{B}^*$, laquelle les synthétise d'une façon efficace vu qu'elle concentre *n* sur une seule modalité ou, au plus, sur deux modalités contiguës et offre donc plus de possibilités interprétatives. Cette distribution a été appelée moyenne bipolaire (*MB*).

Pour obtenir la *MB* d'une distribution de fréquence, il est utile de considérer la fonction:

$$\bar{S} = \frac{1}{n} G = \frac{1}{n}\sum_{s=1}^{k} R_s = \frac{1}{n}\sum_{s=1}^{k} s \cdot n_s = \sum_{s=1}^{k} s \cdot f_s \tag{3}$$

où $f_s = n_s / n$. La (3) met en évidence que $\bar{S}$ correspond à la moyenne arithmétique pondérée de *X*. Les valeurs que $\bar{S}$ prend varient selon la progression arithmétique de raison 1/*n*: 1, 1+1/*n*,…, 2–1/*n*, 2, 2+1/*n*,…, *k*.

La fonction $\bar{S}$ est cohérente avec l'ordre *W*. Sont donc valides, même pour $\bar{S}$, les observations traitées pour *G* relatives aux sous-ensembles en qui est partagé $\mathcal{B}$. La fonction prend les valeurs entières 1, 2,…, *k* lorsque les distributions concentrent leur nombre sur une seule modalité (dans la suite indiquée comme *MB* de 1er type) et les valeurs décimales lorsque les distributions concentrent *n* sur deux modalités contiguës (dans la suite indiquées comme *MB* de 2ème type).

On peut alors déduire la règle opérationnelle qui permet de calculer la *MB* pour toutes les distributions:

i) alors que $\bar{S}$ prend une valeur entière (1, 2,…, *s*,…, *k*) la *MB* concentre *n* sur les modalités respectives 1, 2,…, *s*,…, *k*.

ii) alors que $\bar{S}$ prend une valeur non entière comprise entre *s* et *s*+1, *s*=1, …, (*k*-1), la *MB* concentre sur la modalité *s*+1 la fréquence n_{s+1} égale au produit de la partie décimale de $\bar{S}$ et *n*, et sur la modalité *s* la fréquence $n_s = n - n_{s+1}$.

2.1 L'écart moyen de la moyenne bipolaire

Soit η la *MB* de *X*. Soit N_s la *s*-ième fréquence cumulée croissante de *X*. Soit $\tilde{N}_s$ la *s*-ième fréquence cumulée croissante de η. L'écart moyen de *X* de sa *MB* η est donné par:

$$S_\eta = \frac{1}{n}\sum_{s=1}^{k}\left|N_s - \tilde{N}_s\right| . \qquad (4)$$

Cet indice a été mis en relation avec *l'écart moyen de la moyenne arithmétique* S_μ et on a démontré que:

$$S_\eta \le S_\mu \qquad (5)$$

l'égalité vaut si et seulement si $\mu = s$ pour *s*=1, 2, …, *k*; donc l'égalité entre les deux indices a lieu uniquement dans le cas de la *MB* du 1^er type.

2.2 Distribution de variabilité maximum de *X* et maximum de l'écart moyen de la moyenne bipolaire

Dans un but de confrontation, l'écart moyen de la moyenne bipolaire peut être exprimé comme une mesure relative. On peut faire ça de deux façons: (*i*) en le comparant à une moyenne appropriée; (*ii*) en le comparant à la valeur qu'il prend dans une distribution de variabilité maximum. Dans le cas (*i*) il paraîtrait convenable de choisir la moyenne arithmétique, car on a vu que celle-ci a des liens avec la fonction qui détermine la moyenne bipolaire. Dans le cas (*ii*) il est nécessaire, avant tout, de définir la distribution de variabilité maximum qui est différente selon que le nombre des cas est pair ou impair. Une fois qu'on a défini la distribution appropriée de variabilité maximum, on détermine la moyenne bipolaire correspondant, qui dépend aussi du fait que les modalités du caractère discret (ou les rangs attribués aux catégories du caractère ordinal) peuvent être paires ou impaires.

Distribution de variabilité maximum pour variables discrètes (définition)
Pour distinguer les deux cas de *n* pair ou impair on propose la symbolique suivante: *nombre pair*: *n* (fréquence générique n_s); *nombre impair*: $m = n+1$ (fréquence générique m_s).
Soit *X* une variable discrète qui prend les valeurs 1, 2, …, *s*, …, *k* avec des fréquences dont la somme peut être paire ou impaire.

1^er cas) *Nombre pair*: $n = \sum_{s=1}^{k} n_s$. On définit comme distribution de variabilité maximum de *X*
(1) (*n*/2, 0,…, 0,…, *n*/2).

2^ème cas) *Nombre impair*: $m = n + 1 = \sum_{s=1}^{k} m_s$. On définit comme distributions de variabilité maximum de *X*: (2*a*) (*n*/2+1, 0,…, 0,…, *n*/2) et (2*b*) (*n*/2, 0,…, 0,…, *n*/2+1).

Le maximum de S_η est indiqué avec S_η^*. On le détermine avec la moyenne bipolaire η* représentée par la valeur $\bar{S}*$ que la fonction $\bar{S}$ assume en correspondance de la distribution de variabilité maximum ou, ce qui est le même, par la valeur μ^* que la moyenne arithmétique assume dans une telle distribution [voir (3)]. De la formule (4) de S_η dérive immédiatement l'expression de S_η^*

$$S_\eta^* = \frac{1}{n}\sum_{s=1}^{k}\left|N_s^* - \tilde{N}_s^*\right| \qquad (6)$$

où N_s^* et $\tilde{N}_s^*$ sont les fréquences cumulées qui résultent, respectivement, de la distribution de variabilité maximum et de sa moyenne bipolaire η^*. On a dit que pour la détermination du maximum de S_η^* on doit considérer que *k* aussi peut être pair ou impair. On verra que seulement pour *k* impair et *n* pair la moyenne bipolaire associée à la distribution de variabilité maximum est du 1er type; par contre dans tous les autres cas elle est du 2ème type.

Cas A: k impair et n pair

Si *n* est pair, dans la distribution de variabilité maximum le nombre est équitablement reparti entre la valeur plus petite 1 et la plus grande *k*. Puisque la moyenne arithmétique $\mu^* = (k+1)/2$ est un nombre entier, *k* étant impair, la moyenne bipolaire est du 1er type et concentre *n* sur (*k*+1)/2.

Le maximum de l'écart moyen de la moyenne bipolaire est :

$$S_\eta^* = \frac{k-1}{2}$$

Cas B: k impair et n impair

Si *n* est impair on a deux distributions de variabilité maximum puisque *n* ne peut pas être équitablement reparti entre les valeurs extrêmes, mais la fréquence qui reste doit être attribuée ou à la plus petite (cas *a*) ou à la plus grande (cas *b*). Dans les deux cas la moyenne de la distribution de variabilité maximum n'est pas entière et la moyenne bipolaire est du 2ème type.

Le maximum de l'écart moyen de la moyenne bipolaire est, tant pour le cas *a*) que pour le cas *b*):

$$S_\eta^* = \frac{k-1}{2}\frac{n-1}{n}$$

le résultat est semblable à celui obtenu pour le cas *A* avec *n* pair, ou le terme (*k*-1)/2 est multiplié par le facteur correctif (*n*-1)/*n*.

Cas C: k pair et n pair

Si *n* et *k* sont pairs, la moyenne arithmétique $\mu^* = (k+1)/2$ n'est pas un nombre entier et donc la moyenne bipolaire est du 2ème type.

Le maximum de l'écart moyen de la moyenne bipolaire est:

$$S_\eta^* = \frac{k-2}{2}$$

Cas D: k pair et n impair

Lorsque *k* est pair et *n* impair, la moyenne de la distribution de variabilité maximum n'est pas entière: *n* ne peut pas être équitablement reparti entre les valeurs extrêmes, mais la fréquence qui reste doit être attribuée à une des deux. Les moyennes bipolaires des deux distributions de variabilité maximum sont du 2ème type.

Le maximum de l'écart moyen de la moyenne bipolaire est dans le deux cas:

$$S_\eta^* = \frac{k-2}{2}\frac{n+1}{n}$$

le résultat est semblable à celui obtenu pour le cas *C* avec *n* pair, où le terme (k-2)/2 est multiplié par le facteur correctif $(n+1)/n$.
L'ensemble des résultats obtenus est synthétisé dans le Tableau 2.

	***n* pair**	***n* impair**
***k* impair**	$\frac{k-1}{2}$	$\frac{k-1}{2}\frac{n-1}{n}$
***k* pair**	$\frac{k-2}{2}$	$\frac{k-2}{2}\frac{n+1}{n}$

Tableau 2: Expressions de S_η^* étant donné *n* (nombre d'unités statistiques) et *k* (note maximum)

3. Exemples

Pour illustrer l'emploi de la *MB* dans l'analyse sensorielle, on retourne à la situation présentée au début du travail. On calcule pour les distributions des évaluations de sept juges sur deux descripteurs (*A* et *B*) et pour trois *grappe* (G_1, G_2, G_3) les *MB* selon la règle opérationnelle qu'on a définie.
On présente, comme exemple, le calcule de la *MB* pour G_{1A} et G_{2A}. Dans le cas de G_{1A} étant donné que $\bar{s}$ =1,857 on utilise la règle *ii*) qui attribue la fréquence 7x0,857=6 à la note 2 et la fréquence 7-6=1 à la note 1. Comme on peut voir dans le Tableau 3, on trouve la *MB* (1, 6, 0, 0, 0) qui concentre 86% des fréquences sur la note 2 et 14% sur la note 1. Dans le cas de G_{2A} étant donné que $\bar{s}$ = 2 on utilise la règle *i*) qui attribue toute la fréquence 7 à la note 2. Comme on peut voir dans le Tableau 3, on trouve la *MB* (0, 7, 0, 0, 0) qui concentre 100% des fréquences sur la note 2. De la même façon on trouve la *MB* pour les autres distributions.

Notes	***Descripteur A***			***Descripteur B***		
	***MB*(G_{1A})**	***MB*(G_{2A})**	***MB*(G_{3A})**	***MB*(G_{1B})**	***MB*(G_{2B})**	***MB*(G_{3B})**
1	1	0	0	0	0	0
2	6	7	6	0	0	0
3	0	0	1	7	7	7
4	0	0	0	0	0	0
5	0	0	0	0	0	0
$\bar{s}$ (μ)	**1,857**	**2**	**2,143**	**3**	**3**	**3**
S_η	**0,571**	**0,571**	**0,571**	**1,429**	**0,286**	**1,143**
S_η^*	$\frac{5-1}{2}\times\frac{6}{7}$**=1,714**	**1,714**	**1,714**	**1,714**	**1,714**	**1,714**
$S_\eta/S_\eta^*\%$	**33,31%**	**33,31%**	**33,31%**	**83,37%**	**16,69%**	**66,69%**

Tableau 3: Distributions des évaluations de sept juges pour trois *grappe* sur les descripteurs *A* et *B* et indices correspondants

Les distributions G_{1A}, G_{2A} et G_{3A} appartiennent à trois sous-ensembles différents de $\mathcal{B}$ et sont synthétisées par la *MB* qui se trouve dans le sous-ensemble correspondant. On peut voir que l'ordre de

la *MB* est le même que celui donné par μ, mais la *MB* est une distribution de fréquence qui peut donner lieu à des interprétations plus amples que μ qui fournit comme synthèse une seule valeur.
La situation se présente différemment dans le cas du descripteur *B* dont les trois distributions sont synthétisées par la même *MB*: en effet on a $\bar{s}$ =3 pour G_{1B}, G_{2B} et G_{3B} et si on utilise la règle *i*) qui attribue toute la fréquence 7 à la note 3 on trouve la *MB* (0, 0, 7, 0, 0) qui concentre 100% des fréquences sur la note 3. Il n'est donc pas possible établir un ordre pour les trois distributions qui sont équivalents selon l'ordre *W*.
Dans une situation comme celle-ci, il est utile de considérer les valeurs des écarts moyens de la moyenne bipolaire qui sont différentes pour chacune des trois distributions (Tableau 3). L'indice est 83,37% de son maximum théorique pour G_{1B} , 16,69% pour G_{2B} et 66,69% pour G_{3B}; il est alors possible de classer les jugements sur les trois *grappe* qui sont caractérisées par la même *MB*, en tenant compte de la variabilité des jugements donnés par les sept juges. On sera donc amené à juger meilleure la distribution G_{2B} par rapport à la distribution G_{3B} et celle-ci meilleure par rapport à la distribution G_{1B}.

3.1 Distributions de nombre différent

On considère maintenant des distributions qui viennent des données récoltées à l'occasion de la 35^{éme} édition de *Vinitaly* qui s'est tenue à Vérone en 2001. Au cours de cette manifestation on effectue un test sur le consommateur, appelé *Grappa & C. Tasting.* En général chacun des produits présente une fréquence différente des dégustations, car le nombre de ceux qui se présentent pour déguster l'un au l'autre produit est variable
Parmi les distributions des jugements on a sélectionné celles de cinq *grappe* différentes pour le descripteur goût-olfactif sur une échelle de notation de 1 (*très mauvais*) à 6 (*très bon*). Les distributions (Tableau 4) sont ordonnées selon la valeur croissante de $\bar{s}$ (μ). Étant donné que les fréquences sont de nombre différent, pour la comparaison on fournit aussi les *MB* en pourcentage.

Grappe	$G_1(go)$			$G_2(go)$			$G_3(go)$			$G_4(go)$			$G_5(go)$		
Notes	n_s	η	η%	n_s	η	η%	n_s	η	η%	n_s	η	η%	n_s	η	η%
1	2	0	0	2	0	0	0	0	0	0	0	0	0	0	0
2	2	0	0	2	0	0	1	0	0	1	0	0	0	0	0
3	1	10	100	6	11	73	7	5	31	1	0	0	1	0	0
4	4	0	0	1	4	27	5	11	69	7	4	19	2	0	0
5	1	0	0	3	0	0	2	0	0	4	17	81	10	17	100
6	0	0	0	1	0	0	1	0	0	8	0	0	4	0	0
n	**10**			**15**			**16**			**21**			**17**		
$\bar{s}$ (μ)	**3**			**3,2667**			**3,6875**			**4,8095**			**5**		
S_η	**1,2**			**0,8**			**0,5**			**0,7619**			**0,4706**		
S_η^*	$\frac{6-2}{2}$=**2**			$\frac{6-2}{2}\times\frac{16}{15}$=**2,133**			$\frac{6-2}{2}$=**2**			$\frac{6-2}{2}\times\frac{22}{21}$=**2,095**			$\frac{6-2}{2}\times\frac{18}{17}$=**2,118**		
S_η/S_η^*%	**60%**			**37,51%**			**25%**			**36,36%**			**22,22%**		

Tableau 4: Distributions de fréquence et moyenne bipolaire des notes données par différents dégustateurs pour cinq *grappe* selon le descripteur goût–olfactif et indices correspondants

La distribution $G_1(go)$ est la “plus mauvaise” étant donné qu'elle présent la plus petite valeur de μ (et moyenne bipolaire associée) combinée à la valeur la plus élevée de l'indice de variabilité (60% du maximum). Par contre $G_5(go)$ est la distribution “meilleure”: elle présente la moyenne la plus élevée et le plus petit indice de variabilité (22,22% du maximum). On peut faire des considérations semblables pour les trois autres distributions pour en conclure que l'ordre établi par la moyenne n'est pas modifié par l'indice de variabilité.
On veut maintenant établir ce qui se passerait si, à la place de la distribution $G_1(go)$, on avait observé la distribution substitutive $S_1(go)$ qui est présentée dans le Tableau 5 avec la $G_2(go)$.

Notes	$S_1(go)$		$G_2(go)$	
	n_s	η	n_s	η
1	0	0	2	0
2	1	0	2	0
3	6	8	6	11
4	3	2	1	4
5	0	0	3	0
6	0	0	1	0
n	**10**		**15**	
$\bar{s}$ (μ)	**3,2**		**3,2667**	
S_η	**0,2**		**0,8**	
S_η^*	**2**		**2,133**	
$S_\eta/S_\eta^*\%$	**10%**		**37,51%**	

Tableau 5: Distribution de fréquence $S_1(go)$, substitutive de la $G_1(go)$, et $G_2(go)$

Dans cette situation hypothétique la moyenne de $S_1(go)$ est encore plus petite que celle de $G_2(go)$ mais cette différence est minime (3,2 confronté à 3,2667). Par contre l'indice de variabilité de $S_1(go)$ est dans ce cas sûrement inférieur à celui de $G_2(go)$, vu que la valeur de celle-ci représente 37,51% du maximum théorique, alors que pour la $G_2(go)$ la valeur en représente 10%. Dans ce cas l'ordre établi selon la valeur de μ pourrait être modifié, en jugeant “meilleure” la première distribution; en effet une plus petite variabilité dans les jugements pourrait être plus appréciée, en dépit d'une moyenne légèrement inférieure.

4. Observations conclusives

Dans un travail de 2005, W. Maffenini et Mi. Zenga ont introduit dans la littérature une nouvelle synthèse pour les variables ordinales: la moyenne bipolaire. C'est une distribution qui concentre tous les *n* cas sur une seule des *k* catégories de la variable ou, au plus, sur deux catégories contiguës et qui est cohérente avec le principe de dominance statistique. L'année suivante, Maffenini et Ma. Zenga ont appliqué la moyenne bipolaire aux variables discrètes et ils ont aussi défini pour cette moyenne une mesure de variabilité: l'écart moyen de la moyenne bipolaire. Cette mesure peut être calculée également pour les caractères qualitatifs ordinaux si les catégories sont exprimées sur une échelle d'évaluation en notes.
L'écart moyen de la moyenne bipolaire est un indice absolu de variabilité; comme pour tous les indices de variabilité il est utile de définir sa valeur maximum dans le but de comparaison. On a donc

déterminé (E. Brentari, L. Dancelli, W. Maffenini, 2009) la distribution de variabilité maximum et à partir de cette distribution, on a calculé la moyenne bipolaire correspondante. Ces mesures sont utiles dans l'analyse sensorielle où on se trouve souvent dans la nécessité de comparer les distributions des évaluations exprimées par *n* sujets pour des produits sur une échelle de *k* notes.
Dans cette étude, en ayant recours à différents types de données d'évaluation de produits alimentaires, on a approfondi la capacité de synthèse de la moyenne bipolaire en employant aussi la nouvelle mesure de dispersion. Les distributions utilisées appartiennent à deux groupes qui ont la caractéristique d'avoir le même nombre des cas ou un nombre des cas différents. On se trouve confronté aux cas suivants:

- distributions qui ont la même moyenne arithmétique et, par conséquent, la même moyenne bipolaire, selon l'ordre induit par ces synthèses, sont considérées équivalentes. Si les mesures respectives de variabilité prennent une valeur sensiblement différente, on peut dépasser l'indécision en reconnaissant à la distribution moins variable une supériorité par rapport à l'autre, vu qu'elle met en évidence une moindre "dispersion" des jugements;
- distributions avec des moyennes très voisines peuvent être ordonnées d'une façon différente de ce qu'elles le seraient sur base de leur moyenne bipolaire, si les mesures de variabilité qui leur sont associées sont assez différentes.

Bibliographie

Brentari E., Dancelli L. (2005), Sull'impiego della media aritmetica nell'analisi sensoriale, Rapporti di Ricerca del Dipartimento Metodi Quantitativi, Università degli Studi di Brescia, n. 260, pp. 1-24.

Brentari E., Dancelli L. (2008), Media, Mediana e Media Bipolare: semplici strumenti per confrontare i prodotti, *L'assaggio*, n. 24, pp. 13-16.

Brentari E., Dancelli L., Maffenini W. (2009), Osservazioni sullo scostamento medio dalla media bipolare, Rapporti di Ricerca del Dipartimento Metodi Quantitativi, Università degli Studi di Brescia, n. 338, pp. 1-27.

Maffenini W., Zenga Mi. (2005), Bipolar Mean for ordinal variables, *Statistica & Applicazioni*, III, (1), pp. 3-18.

Maffenini W., Zenga Ma. (2006), Bipolar Mean and Mean Deviation about the Bipolar Mean, *Statistica & Applicazioni*, IV, (1), pp. 35-53.