

Considérations sur l'analyse statistique de données sensorielles

Jean-Jacques Claustrioux

Unité de Statistique et Informatique. Faculté universitaire des Sciences agronomiques de Gembloux. Passage des Déportés, 2. B-5030 Gembloux (Belgique).

Reçu le 16 janvier 2001, accepté le 10 juillet 2001.

Cet article introduit les principaux éléments auxquels il convient d'être attentif pour définir, mesurer et analyser des caractéristiques perçues par l'intermédiaire des organes des sens.

Mots-clés. Analyse sensorielle, épreuve sensorielle, évaluation sensorielle, statistique sensorielle.

Considerations about statistical analysis of sensory data. This publication introduces some steps to build an adequate statistical data analysis for sensory studies.

Keywords. Sensory analysis, sensory assay, sensory assessment, sensory statistics.

1. INTRODUCTION

L'analyse sensorielle ou évaluation sensorielle permet de définir, mesurer, analyser et interpréter les caractéristiques d'un produit perçues par l'intermédiaire des organes des sens, c'est-à-dire ses propriétés gustatives, olfactives, visuelles, auditives et tactiles. Certaines normes définissent simplement l'analyse sensorielle comme suit : examen des propriétés organoleptiques d'un produit par les organes des sens.

Dans ce cadre, l'homme, appelé juge ou répondant, est considéré comme instrument de mesure chaque fois qu'il n'existe pas de capteur physique capable de rivaliser avec son équivalent sensoriel (Mac Leod, Sauvageot, 1986), c'est-à-dire lorsque les méthodes instrumentales ne permettent pas de décrire et de quantifier les caractéristiques d'un produit telles que l'homme les perçoit.

En principe, pour tout domaine qui relève davantage des techniques expérimentales, on ne réalise pas l'analyse statistique des données sans avoir préalablement réfléchi à l'étape de conception des essais sensoriels ou épreuves sensorielles. Il est indispensable d'insister sur le fait que, "derrière" cette phase conceptuelle qui précise également toute la partie relative à la mise en œuvre proprement dite des épreuves, sont cachés des modèles statistiques fondamentaux utilisés pour analyser les données.

C'est pourquoi, après cette introduction, le traitement statistique des données sensorielles est évoqué en étroite relation avec la description des différents types d'épreuves sensorielles. Indirectement, elles

précisent les objectifs poursuivis. Elles sont groupées en deux grandes catégories, à savoir, d'une part, les épreuves analytiques et, d'autre part, les épreuves hédoniques.

Il est aussi nécessaire d'évoquer des éléments relatifs à l'organisation proprement dite de ces épreuves sensorielles ou choix d'un dispositif expérimental adapté, avant d'envisager quelques considérations sur le traitement statistique proprement dit des données sensorielles, y compris des réflexions générales sur les logiciels utilisés. Enfin, quelques conclusions terminent le document.

2. ÉPREUVES ANALYTIQUES

Les épreuves analytiques sont destinées à mettre en évidence des différences entre produits ou à décrire les propriétés sensorielles des produits, abstraction faite du sentiment de satisfaction ou au contraire de déplaisir qu'ils peuvent procurer (Delvaux, 1992).

Les épreuves analytiques se subdivisent en épreuves discriminatives, épreuves de similitudes, épreuves de classements, épreuves d'intervalles et épreuves de descriptions quantifiées.

1) Les épreuves discriminatives (*discriminative tests*, *difference tests*) visent à détecter la présence ou l'absence de différences sensorielles globales entre des produits, la caractéristique sensorielle sur laquelle portent les éventuelles différences n'étant pas toujours connue des juges *a priori*. L'épreuve A-non A,

l'épreuve par paires, l'épreuve duo-trio, l'épreuve triangulaire, etc., sont quelques exemples d'épreuves de cette catégorie. Les variables observées sont souvent alternatives et leur traitement statistique est assez élémentaire.

2) Les épreuves de similitudes (*similarity tests*) consistent à déterminer globalement l'importance des différences perçues entre les échantillons de plusieurs produits, souvent deux produits, en affectant à chaque couple de produits une distance reflétant au mieux les différences sensorielles globales perçues entre deux échantillons et en notant sur une échelle structurée ou non l'intensité de cette différence.

Pour une échelle structurée, la signification de chaque niveau de l'échelle est indiquée au juge de façon discrète (exemple : très sec, sec, ni sec–ni juteux, juteux, très juteux) ou continue (exemple : échelle variant de un à neuf et attribuant la valeur un à la note très dur, trois à la note dur, cinq à la note ni dur–ni tendre, sept à la note tendre et neuf à la note très tendre).

Pour une échelle non structurée, seules les deux limites extrêmes sont définies (exemple : échelle d'une longueur de dix centimètres portant aux extrémités les indications "échantillon ne présentant pas d'odeur de menthe" et "échantillon présentant une odeur de menthe").

La définition de ces types d'échelles permet déjà d'appréhender quelques difficultés pour le traitement statistique et surtout pour les conclusions statistiques associées aux données sensorielles. Combien de notes faut-il définir ? Si par rapport à l'origine, on mesure en unité métrique le degré de similitude entre deux produits, quelle longueur doit avoir l'échelle de référence ? Que signifie l'écart entre deux notes ? Quatre est-il le double ou la moitié de deux : deux fois plus tendre ou deux fois moins sec (intervalle variable ou constant) ?

3) Les épreuves de classements ou épreuves de rangements (*ranking tests*) consistent à ranger deux ou plusieurs échantillons de produits selon l'intensité croissante ou décroissante d'une caractéristique sensorielle.

Outre les difficultés de définition du paramètre à apprécier, s'ajoute le fait qu'il est bien difficile à un seul juge de comparer simultanément un grand nombre de produits, ce dernier dépassant rarement en pratique sept produits.

4) Les épreuves d'intervalles, de notations ou de cotations (*rating tests, scoring tests*) concernent l'évaluation de l'intensité d'une ou de plusieurs caractéristiques sensorielles pour plusieurs échantillons de produits, sur base des échelles décrites précédemment en vue de quantifier des différences, en

considérant *a priori* pour l'analyse statistique chaque caractère indépendamment l'un de l'autre.

5) Les épreuves de descriptions quantifiées aussi appelées analyses descriptives quantitatives (*quantitative descriptive analyses*) visent à établir un profil sensoriel complet d'un ou de plusieurs produits.

Une épreuve de ce type se décompose en trois étapes (Barthélémy *et al.*, 1990), à savoir la recherche des caractéristiques sensorielles les plus pertinentes pour exprimer un maximum d'informations sur les propriétés sensorielles des produits étudiés, la mesure de l'intensité de la sensation perçue pour chaque descripteur choisi et la représentation visuelle du profil des produits par des pseudo-histogrammes, des graphes segmentés ou des graphiques polaires.

3. ÉPREUVES HÉDONIQUES

Un produit alimentaire d'apparence laide n'a aucune chance d'être commercialisé, même si ses qualités gustatives sont exceptionnelles. Les épreuves hédoniques doivent tenir compte de ce fait ; elles concernent l'étude des préférences et des aversions des consommateurs, des utilisateurs ou des clients (Delvaux, 1992).

On distingue des épreuves hédoniques de classements et d'évaluations y compris des épreuves de lassitude, d'aversion, d'authenticité et du profil du produit idéal.

1) Les épreuves de classements ont les mêmes caractéristiques que celles décrites précédemment mais avec des qualificatifs plus subtils comme le caractère agréable ou non des produits (de moins agréable à plus agréable).

Il est évident que le traitement statistique concerne davantage le niveau relatif de préférence et non le niveau absolu du caractère apprécié, c'est-à-dire l'analyse de variables particulières que sont les proportions.

2) Quant aux épreuves d'évaluations, elles sont semblables aux épreuves d'intervalles décrites précédemment.

4. DISPOSITIFS EXPÉRIMENTAUX

1) Les épreuves de classements et d'intervalles étant fréquemment employées, il est intéressant de s'attarder à quelques brèves considérations sur l'allocation des objets, par exemple des produits, aux juges.

En pratique, au niveau du protocole expérimental, les épreuves de ces types ne concernent que l'étude des objets d'un seul facteur fixe, comme par exemple

différents fournisseurs, des types de morceaux de viande, différentes boissons, des bains mousses, etc. C'est pourquoi, on envisage essentiellement des dispositifs expérimentaux adaptés aux expériences non factorielles.

Parmi tous les protocoles expérimentaux disponibles, il faut citer le dispositif classique en blocs aléatoires complets où chaque juge reçoit tous les produits selon un ordre aléatoire, les différents dispositifs en blocs aléatoires incomplets, principalement équilibrés, lorsque chaque juge ne peut pas nécessairement déguster tous les produits et le carré latin ou le *cross-over* pour faire déguster par chaque juge chaque produit un même nombre de fois en premier lieu, en deuxième lieu, en troisième lieu, etc., au cours de plusieurs périodes.

2) Des études plus récentes sur l'hétérogénéité et l'effet de compétition dans des essais comparatifs de variétés de plein champ (Claustriau, 1997) ont abouti à la mise au point de nouveaux dispositifs qui peuvent s'avérer être très utiles pour la conception des épreuves sensorielles, à savoir les dispositifs équilibrés pour le voisinage. Ces dispositifs complets permettent de faire en sorte que chaque produit soit dégusté par chaque juge un même nombre de fois après chacun des autres produits (Azaïs *et al.*, 1993), certes en augmentant la durée totale d'évaluation pour chaque juge.

À titre d'exemple, la **figure 1** présente ce type de plan pour l'étude de sept produits par six juges (lignes) ; abstraction faite de la première colonne et de la dernière colonne, chaque produit est précédé et suivi de chacun des six autres objets une seule fois.

1	6	5	2	7	3	4	1	6
1	3	2	6	4	7	5	1	3
7	4	6	2	3	1	5	7	4
1	4	3	7	2	5	6	1	4
5	4	2	1	7	6	3	5	4
4	5	3	6	7	1	2	4	5

Figure 1. Exemple de dispositif équilibré pour le voisinage
— *Example of neighbour-design.*

5. ANALYSE STATISTIQUE

1) Les méthodes statistiques fréquemment utilisées pour analyser des données sensorielles et présentées en fonction du type d'épreuve ont déjà fait l'objet de publications antérieures (Claustriau, Delvaux, 1992). C'est pourquoi, on insiste davantage sur quelques particularités de ce type de données.

Selon le nombre de variables ou caractéristiques observées simultanément, les outils statistiques

disponibles se classent en méthodes à une dimension ou à plusieurs dimensions. De plus, selon que les méthodes sont contraignantes ou peu contraignantes, principalement quant à la condition d'application relative aux distributions des populations-parents, elles sont qualifiées de paramétriques ou non paramétriques (*distribution-free methods*).

Souvent, dans le cas de classements hédoniques, ce sont ces méthodes qui sont employées puisque les données correspondent à des rangs.

2) À titre d'exemple de traitement de données de ce type et afin de montrer que le recours à des outils sophistiqués n'est pas toujours nécessaire, le **tableau 1** présente les résultats du classement par sept juges de viandes provenant de quatre fournisseurs, tous considérés *a priori* comme équivalents, sur base de l'appréciation de la saveur de la viande (Bastiaens, Delvaux, 1992).

Le test de Friedman (Dagnelie, 1998) relatif à l'hypothèse nulle d'égalité des barycentres ou rangs moyens conduit au rejet de l'hypothèse ($\alpha = 0,05$), la valeur de Pearson valant 9,86 avec trois degrés de liberté.

Dès lors, au niveau de signification considéré, les écarts entre les barycentres ne peuvent donc pas être imputés simplement au hasard. La structuration des barycentres par l'intermédiaire d'une méthode comparable à la technique classique de comparaisons multiples des moyennes aboutit à des conclusions assez complexes : $A=B$ et $B=C=D$.

Certes, des méthodes plus élaborées d'analyse à plusieurs variables, en particulier la classification ascendante hiérarchique, pourraient être utilisées pour compléter ces premières conclusions en évaluant l'appréciation des juges, technique aussi appelée segmentation ou typologie des juges.

Une nouvelle présentation simple des résultats permet facilement de vérifier si les premières conclusions, notamment la seconde ($B=C=D$), correspondent à un jugement analogue ou consensuel des juges. Le **tableau 2** reprend le nombre d'occurrences des différents rangs par fournisseur, calculé sur l'ensemble des juges.

Tableau 1. Exemple de résultats d'un classement sensoriel
— *Example of results of a sensory ranking test.*

Fournis- seurs	Juges							Bary- centres
	1	2	3	4	5	6	7	
A	1	2	2	1	1	1	1	1,3
B	2	1	1	4	3	2	4	2,4
C	3	3	4	2	4	4	3	3,3
D	4	4	3	3	2	3	2	3,0

Tableau 2. Exemple de la distribution de fréquences des rangs attribués à chaque fournisseur par l'ensemble des juges — *Example of frequency distribution of ranks given by the judges of the sensory panel to each assessed item (here several types of deliverers).*

Fournisseurs	Rangs			
	1	2	3	4
A	5	2	0	0
B	2	2	1	2
C	0	1	3	3
D	0	2	3	2

Cette technique simple montre clairement une distinction entre le fournisseur B et les fournisseurs C et D, tenant compte de l'appréciation différente des juges.

3) Enfin, en matière de traitement informatique des résultats, la plupart des logiciels statistiques classiques permettent de traiter des données sensorielles, pour autant que la méthode d'analyse soit préalablement bien définie, afin de choisir judicieusement la ou les procédures adéquates.

Des logiciels statistiques spécifiques à l'analyse sensorielle sont disponibles. Ils reprennent les procédures les plus fréquemment utilisées dans le domaine. Leur coût est à mettre en rapport, notamment, avec l'avantage qu'ils offrent de guider l'utilisateur pour le choix du test statistique, en principe adapté directement à la situation particulière, voire parfois à l'enfermer dans une voie sans issue. Il faut cependant ajouter que ce type d'outils spécifiques intègre aussi toute la gestion des séances de dégustation (création des séances, établissement des questionnaires, déroulement des séances et acquisition des données).

6. CONCLUSIONS

La phase de conception d'une épreuve sensorielle est l'élément clef de toute analyse sensorielle.

Elle conditionne dès le départ le bon usage par la suite des méthodes statistiques, avec ou sans logiciel, et, par conséquent, elle détermine déjà dès le départ la qualité des conclusions d'une étude.

Si besoin, il ne faut jamais hésiter à consulter le statisticien, qui doit toujours être perçu, non seulement, comme un spécialiste du traitement des données, mais surtout, comme celui qui peut aussi apporter une assistance déterminante en matière de planification d'épreuves sensorielles.

Bibliographie

Azaïs JM., Bailey RA., Monod H. (1993). A catalogue of efficient neighbour-designs with border plots. *Biometrics* **49**, p. 1252–1261.

Barthélémy J., Clément JF., Danzart M., Issanchou S., Köster EP., Mac Leod P., Nicod H., Sauvageot F., Sztrygler F., Touraille C. (1990). *Évaluation sensorielle. Manuel méthodologique*. Paris : Lavoisier, 328 p.

Bastiaens A., Delvaux A. (1992). Exemples d'analyses sensorielles. *Ann. Gembloux* **98**, p. 129–139.

Claustriaux JJ. (1997). Hétérogénéité et compétition. In *Les effets de compétition dans les essais variétaux de plein-champ*. France, Boigneville : ITCF, Paris : INRA, p. 9/96– 15/96.

Claustriaux JJ., Delvaux A. (1992). Traitement statistique de données sensorielles. *Ann. Gembloux* **98**, p. 117–128.

Dagnelie P. (1998). *Statistique théorique et appliquée. Tome 2 : Inférence statistique à une et à deux dimensions*. Bruxelles : De Boeck, 659 p.

Delvaux A. (1992). Les épreuves sensorielles. *Ann. Gembloux* **98**, p. 105–115.

Mac Leod P., Sauvageot F. (1986). Bases neurophysiologiques de l'évaluation sensorielle des produits alimentaires. *Les Cahiers de l'Ensbana* **5**. Paris : Lavoisier, 165 p.

(8 réf.)