

Analyse sensorielle (à suivre le cours avec le polycopié qui présente des tables, des figures....)

I) Introduction

L'analyse sensorielle (AS) est l'ensemble des méthodes permettant d'évaluer les qualités organoleptiques d'un produit. Cette analyse fait appel aux organes des sens (vue, odorat, touché, ouïe...).

L'analyse sensorielle est née dans les années 60's aux USA et est devenue indispensable en complément des analyses physico-chimiques et de l'évaluation d'un produit.

Domaines d'utilisation de l'AS :

On utilise l'analyse sensorielle en R&D : outil de développement et de caractérisation.

En R&D : Utilisée soit

- en formulation (produits nouveaux, imitation de produits concurrents),
- soit pour la comparaison
 - o évaluation des produits de la concurrence (matières premières (MP), recette, procédé...)
- pour l'interprétation, corrélation avec des données instrumentales
 - o couleurs/colorimètre ; texture/texturométrie ; CGG/olfactométrie

En contrôle qualité (ou production), l'analyse sensorielle est l'outil d'évaluation et de calibration. On l'utilise en début de chaîne (sélection de MP), en cours de chaîne de production, en contrôle de produits finis en comparant avec le standard. On peut l'utiliser pour mettre en évidence l'effet des conditions de stockage sur les qualités sensorielles des produits.

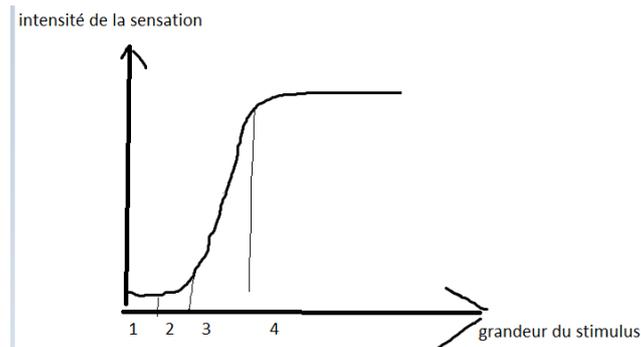
Qui utilise l'AS :

L'analyse sensorielle est utilisée en R&D, en industrie agroalimentaire, cosmétique, pharmaceutique, automobile (odeur de voiture neuve), électroménager, textile...

II) Perception sensorielle

1) Relation sensation-stimulus

L'intensité de la sensation perçue (olfactive, gustative...) est dépendante de la grandeur de stimulus (*agent capable de provoquer la réaction de système excitable (récepteurs)*). (voir figure ci-dessous)



La zone 1 = zone infra-luminaire = c'est une zone dans laquelle la sensation est confuse (bruit de fond, car la concentration en molécule est faible, zone en zigzag).

La zone 2 = zone lumineuse = la sensation est perçue de façon aléatoire, l'intensité étant faible. C'est dans cette zone qu'on détermine le seuil de perception/détection d'un composé.

La zone 3 = zone supra-luminaire = La sensation est nette, l'intensité variant de faible à forte quand la concentration de la molécule (=grandeur de stimulus) augmente. C'est dans cette zone qu'on détermine le seuil de reconnaissance d'un composé.

La zone 4 = zone de saturation = la sensation ne varie plus en intensité quand la concentration de la molécule augmente. Ceci peut s'accompagner par une gêne (maux de tête, douleur...).

2) Caractéristiques d'un instrument de mesure

En analyse sensorielle, les tests sont réalisés par des personnes appelées sujets.

On va chercher chez les « sujets » (ou dégustateur des caractéristiques de mesures

- fiabilité/fidèles : aptitude à donner des mesures très voisines lors de l'application du même signal d'entrée
- répétabilité : réponses fidèles quand l'analyse sensorielle est faite dans les mêmes conditions (même salle, mêmes tests...)
- reproductibilité : répétition faite en faisant varier certaines conditions d'une répétition à l'autre (lieux, résultats inter laboratoires...)
- justesse : moyenne des indications données et très voisine de la valeur vraie de la grandeur à mesurer

3) Dispersion des résultats

La dispersion des résultats est due soit :

- A la dispersion due aux produits
 - o Pour réduire la variance due aux produits, on tâchera de présenter au sujet des produits homogènes ou des produits représentatifs
 - Exemple : comparaison de boîtes de conserve de deux usines, on prend des boîtes de lots différents
- A la dispersion due aux sujets

- liée à la variance interindividuelle, dû aux différences de réponses d'un sujet à l'autre.
 - Ceci peut être acceptable si les sujets sont bien entraînés sur les produits à analyser.
 - Dans le cas contraire, on va entraîner les sujets et reconduire les tests d'analyse sensorielle.
- Variance intra-individuelle
 - Ceci est dû aux différences d'une réponse à l'autre du même sujet et doit être stabilisé à sa valeur la plus faible possible grâce aux séances d'entraînement.

En faisant une analyse de variance, on peut connaître la part de chacune de ces dispersions (on le verra plus tard).

III) Formation de groupes d'analyse sensorielle

Les sujets requis pour effectuer des évaluations sensorielles forment des ensembles ou sous-ensembles qu'on appelle groupes.

Le thème panel doit être réservé à un groupe sur les périodes successives répond aux mêmes questions et réalise les même tâches (ex : évolution de l'arôme d'un vin au cours de stockage).

On peut classer les groupes d'évaluation sensorielle :

- Groupes à vocation hédonique
 - Le sujet exprime simplement son avis concernant le caractère agréable que lui confère le produit (j'aime/j'aime pas)
- Groupes à vocation qualitatif et quantitatif
 - Doit avant tout être objectif
 - Qualités de ce groupe = celles qu'on souhaite trouver dans un appareil de mesure

On trouve différents types de sujets dans ce groupe :

- Sujets naïfs
 - Ils ne doivent répondre à aucun critère particulier

Avec eux, on conduit des tests consommateurs
- Sujets qualifiés
 - Ils ont été sélectionnés et entraînés
- Sujet experts
 - Ils ont été entraînés à l'utilisation de méthodes d'AS et sont capables d'effectuer de façon fiable l'analyse sensorielle des produits (Afnor, 1993).

La constitution de groupes d'AS est une tâche importante et doit être réalisée avec soin car de ces réponses dépendent les résultats des analyses. La constitution de groupes d'AS comporte généralement plusieurs étapes :

- Le recrutement et sélection préliminaire des sujets
 - o Il peut y avoir des questionnaires
 - o Le recrutement peut se faire en interne ou externe à l'entreprise, tout dépend de l'objectif de l'AS.
- La sélection des sujets
 - o La première souvent se fait par un entretien (convocation), l'organisateur analyse si
 - le sujet est intéressé,
 - s'il connaît le produit,
 - s'il a un bon vocabulaire pour le produit considéré (bien décrire)
- L'entraînement général des sujets
 - o familiarisation des sujets au vocabulaire
 - o mémorisation des goûts (arômes, saveurs...)
- Le contrôle des sujets
 - o vérification de la fiabilité et répétabilité des réponses (tests statistiques)

IV) Epreuves/tests

IV.1. Choix de l'épreuve

Le choix de l'épreuve se fait en fonction du problème à résoudre et des produits.

Voir page 2 (P2) du polycopié : Les épreuves préliminaires par l'expérimentateur à l'aide de 2 à 3 personnes permettent de choisir deux types d'épreuves :

- Discriminatives (si les différences sensorielles entre les produits sont difficiles à mettre en évidence)
- Descriptives (permet de connaître la nature et la grandeur des différences sensorielles)

IV.2. Organisation pratique des épreuves

Préparation des échantillons : Les échantillons sont préparés dans des récipients identiques, propres, adéquats, à une certaine température.

Ce qui est important, c'est la codification des échantillons. On doit respecter l'anonymat. Il ne doit pas y avoir une impression de gradation. Cette codification se fait utilisant un nombre à 3 chiffres tirés au hasard (ex : 329). Elle doit respecter l'anonymat. Les échantillons sont présentés dans un ordre également tiré au hasard pour éviter des biais.

Le nombre d'échantillons doit être limité et dépend de type d'évaluation sensorielle. Il faut commencer par les produits les moins agressifs (ex : vin blanc, puis vin rouge).

Le choix des sujets (naïfs, experts, qualifiés...) se fait en fonction des tests à conduire.

Les tests sont réalisés en fonction des produits, mais souvent vers 10-11h ou vers 17h. Avant le test, on présente souvent l'objectif du test.

La salle d'analyse sensorielle doit être dans les normes AFNOR (page 3 du PC), loin des bruits, inodore. Dans cette salle, il y a :

- local de préparation
 - o préparation, codage des échantillons
- zone de distribution
 - o présentation des produits
- cabines d'essai
 - o isolements avec siège, pas de communication entre les sujets
 - o possibilité de rinçage de bouche
 - o luminosité
 - o température 20°C (ambiante)
 - o humidité peu élevée (70-80%)
- bureau
- salle de réunion
 - o présenter les résultats au panel

P4 du PC : Les sujets reçoivent une fiche (de dégustation ou d'analyse sensorielle) bien rédigée, claire.

2 logiciels dédiés à l'analyse sensorielle sont utilisés, Fizz et Tastel. Ils permettent de préparer les tests (codification, ordre de distribution par sujets), de réaliser la saisie des données, d'analyser les données...

IV.3. Notions d'hypothèse

En AS, les tests ont pour but de vérifier une hypothèse : celle-ci être exposés clairement et sans ambiguïté, de telle façon qu'elle puisse être traduite en langage statistique sous forme d'une hypothèse nulle H_0 .

H_0 indique qu'il n'existe pas de différence entre les produits soumis à l'essai (par exemple, ces deux essences possèdent la même odeur). C'est souvent l'hypothèse de départ.

L'hypothèse alternative H_1 , est celle qui sera acceptée en cas de rejet de l'hypothèse nulle (ex : ces deux essences n'ont pas la même odeur).

Lorsqu'on conclut sur les tests, il y a deux risques :

- Rejeter une hypothèse juste
- Accepter une hypothèse fautive

Réalité	Décision	
	H_0	H_1
H_0	$1-\alpha$	α
H_1	β	$1-\beta$

- Risque α = erreur de première espèce, en contrôle qualité appelé risque de producteur
 - o = risque de se tromper quand H_0 est rejeté (doit être le plus faible possible). Dans la pratique, le risque le plus élevé qu'on prend est : $\alpha = 0,05$ ou 5%.

- Déterminé sans ambiguïté
- Risque β = risque de 2^e espèce, en contrôle qualité appelé « risque de consommateur ».
 - probabilité de décider que l'hypothèse est juste alors qu'elle est fausse
 - très dépendant de la capacité du groupe à générer de bonnes réponses. Il est déterminé dans des tests préliminaires. Les tables statistiques proposent le risque β à 50, 37.5 et 25% (P8-11 du PC)
 - intéressant mais peu utilisé car montre la puissance du test (un exemple plus tard).

IV.4. Epreuves discriminatives

Elles visent à détecter la présence/absence de différences sensorielles entre deux produits. Elles s'appliquent chaque fois que l'on veut comparer deux ou plusieurs produits entre lesquels les différences sont inconnues et faibles. Mais elles ne permettent ni d'identifier, ni de quantifier les différences.

Ce sont les épreuves les plus utilisées. On se situe dans le domaine lumineuse, donc au niveau du seuil de détection.

Elles sont utilisées pour :

- contrôler la constance d'une production,
- tester une nouvelle technologie de fabrication
- et étudier les conséquences d'un changement de MP, de vieillissement...

A noter que pour chaque épreuve, il existe des normes Afnor. Ici les épreuves principales sont traitées.

IV.4.1. Epreuve triangulaire (NF-V09-103)

Il consiste à présenter deux produits (A et B), dont l'un est répété. Le sujet reçoit donc 3 échantillons (2A 1B ou 1A 2B) codés et il doit indiquer quel échantillon est différent (celui présenté une seule fois).

On a donc 6 possibilités de présentation :

- AAB
- ABA
- BAA
- BBA
- BAB
- ABB

Probabilité de bonne réponse $p = 1/3$: On a une chance sur 3 de donner la bonne réponse.

La probabilité d'échec $q = 2/3 = (1 - p)$.

C'est une épreuve simple à réaliser, utilisée pour déceler des différences entre les produits. A éviter lorsque les tests se succèdent (fatigant). Ce sont des tests à choix forcé (obligation de donner une réponse).

C'est un test utilisable pour la détermination du seuil de perception. A ne pas utiliser pour produits agressifs.

Exemple : un fabricant de soda veut changer de fournisseur pour ses extraits aromatiques. Il élabore un nouveau produit et le compare à son ancien produit lors d'une épreuve triangulaire.

- Nombre de sujets $n = 32$
 - o 17 bonnes réponses
 - o 15 mauvaises
- Peut-on conclure que les deux produits sont perçus comme différents ?

H_0 = les deux produits ne sont pas différents (hypothèse de départ)

Conclusion du test : voir P5, table 2 (établie à partir de la loi binomiale). Cette table donne le nombre de réponses correctes pour conclure qu'il y ait une différence entre les produits et ce, à partir d'un risque α 0.05, le risque α maximum que l'on prenne en AS. Probability levels dans cette table correspond aux risques α .

- On peut dire que les échantillons sont différents avec un risque α de 0,02 (ou 2%), car 17 bonnes réponses avec 32 sujets.
- On rejette donc l'hypothèse de départ (H_0). On accepte l'hypothèse alternative (H_1) en concluant que les deux produits sont différents avec un risque alpha de 2%, c.a.d. avec un risque de se tromper de 2%.

Remarque : la table de la page 7 est établie à partir de la table page 5 pour montrer les risques alpha les plus utilisés dans la pratique.

* risque $\alpha=0,05$

** risque $\alpha=0,01$

** risque $\alpha=0,001$

A noter que la 2^{ème} colonne concerne le test duo-trio (que l'on verra par la suite), la 3^{ème} colonne concerne le test triangulaire.

Dans l'exemple donné ci-dessus on rejette H_0 et on accepte H_1 . Par conséquent on prend un risque β . Il vaut mieux que ce risque β soit le plus faible possible pour que la conclusion du test triangulaire soit le plus plausible possible.

Détermination du risque β dans notre exemple (32 sujets, 17 bonnes réponses) :

Voir la table (ou appendix 1) pour le test triangulaire (page 8) :

n correspond au nombre de réponses (ou sujets) : ici 32

x correspond au nombre de bonnes réponses : ici 17

α correspond au risque α

Concernant le risque β , on observe trois colonnes : 50%, 37,5 % et 25 %.

Extrait de la table (appendix), p 8, pour notre exemple :

			β		
n	x	α	50%	37,5%	25%
32	17	2	4	22	

Avec cette table on constate que les produits ont été trouvés différents avec un risque α de 0,02 (ou 2%), comme cela a déjà été constaté. Quant au risque β , il est de 4% (ou 0,04) avec un groupe capable de donner de bonnes réponses à hauteur de 50%, alors que ce risque est de 22 % avec un groupe qui génère de bonnes réponses à hauteur de 37,5 %. Cela veut dire que les résultats du test sont plus puissants avec un groupe capable de donner de bonnes réponses à hauteur de 50%. En effet lorsque ce groupe décide que les produits sont différents, le risque d'erreur dans leur conclusion ou le risque de se tromper est de 4 % (ou encore 96 % de chance que leur conclusion est juste), alors qu'il est de 22 % dans l'autre cas. Par conséquent, il est intéressant de connaître la capacité du groupe à donner de bonnes réponses. Ceci est déterminé dans des tests préliminaires.

Remarque : les tables des pages 8 à 11 donnent les risques α et β pour les tests triangulaires et duo-trio.

L'exercice ci-dessus sur le soda peut être également analysé par un test dit Chi² (Khi-deux) ou x² pour savoir si les 2 produits sont différents ou non

Rappel de l'exercice : un fabricant de soda veut changer de fournisseur pour ses extraits aromatiques. Il élabore un nouveau produit et le compare à son ancien produit lors d'une épreuve triangulaire.

- Nombre de sujets n = 32
 - o 17 bonnes réponses
 - o 15 mauvaises
- ➔ Peut-on conclure que les deux produits sont perçus comme différents ?

On calcule le Chi² observé (avec les résultats du test).

$$\text{Chi}^2 \text{ obs} = \sum (n_i - e_i)^2 / e_i$$

- n_i = probabilité observée
- e_i = probabilité théorique
- ➔ Si Chi² observé \geq Chi² théorique (avec un risque α et ddl), on conclut que les produits ont été trouvés différents. Dans le cas contraire, on conclut que les produits sont identiques. (rappel : le risque α le plus élevé que l'on prenne en AS est de 0.05).

Pour notre exemple,

- Probabilité théorique :
 - o Réponse correcte = $32 \times (1/3) = 10,7$
 - o Réponse incorrecte = $32 \times (2/3) = 21,3$
- $$\text{Chi}^2 \text{ obs} = (17 - 10,7)^2/10,7 + (15 - 21,3)^2/21,3 = 5,6$$

Pour trouver Chi^2 théorique, voir la table χ^2 de la page 14

On prend comme ddl = (produits - 1) = 2 - 1 = 1

Conclusion = On rejette H_0 , on conclut que les deux produits sont différents avec un risque alpha de 2,5%, car $5,02 < 5,6$.

IV.4.2. Epreuve duo-trio

Cette épreuve consiste à présenter deux produits (A et B), dont l'un est présenté comme témoin en premier. Les deux autres étant présentés au hasard.

La *question* posée : Lequel des deux échantillons est identique au témoin ?

- A puis A et B
 - A puis B et A
 - B puis A et B
 - B puis B et A
- Probabilité de réussite $p = 1/2$, on facilite la tâche de l'expérimentateur comparé au test triangulaire

Ce test est utilisé pour les produits agressifs et quand on dispose d'un échantillonnage limité.

Le résultat est exploité par la loi binomiale, grâce à des tables (pages 6, 7, 10 et 11).

Exemple : 30 sujets réalisent une épreuve duo-trio. 19 réponses correctes, 11 réponses incorrectes.

Pour conclure voir la table de la page 7 : les produits ne sont pas différentes car il fallait 20 réponses correctes pour conclure la différence avec un risque de 0,05 (*). Notons encore que le risque le plus élevé utilisé en AS est de 0,05.

IV.4 .3. Epreuves p parmi n

On compare toujours 2 produits mais avec un nombre d'échantillons plus important pour diminuer la probabilité de donner la bonne réponse au hasard.

Plusieurs combinaisons sont envisageables : 1 parmi 4, 2 parmi 5 et 3 parmi 7.

Le test 2 parmi 5 est le plus utilisé. Un jeu de 5 échantillons codés provenant de 2 produits, dont 2 sont issus d'un produit et 3 de l'autre produit. Le sujet doit regrouper d'une part les 2 échantillons identiques et d'autre part le groupe de 3 échantillons identiques. Ce test est plus

difficile que le test triangulaire et le duo-trio. La probabilité de donner la bonne réponse au hasard est de 1 /10. A éviter sur les produits agressifs et sur plusieurs séries de test. Ce test est intéressant si le nombre de sujet est faible (ex : 5 personnes). Les résultats sont évalués en utilisant la table de la page 15.

IV.5. Épreuve de classement

Ce test consiste à ranger par ordre croissant ou décroissant des échantillons qui sont présentés simultanément.

Il présente 3 caractéristiques :

- Le sujet évalue tous les échantillons avant de donner sa réponse
- Ce test ne renseigne pas sur la position absolue des échantillons sur une échelle d'intensité
 - o On classe, mais on ne donne pas de valeurs d'intensité.
- Ce test ne donne que des indications indirectes des différences entre les échantillons

Ce test est facile à mettre en œuvre, car ne demande pas de produits de référence. Il est facile à comprendre, même par des sujets peu entraînés (voir page 16, un exemple de fiche de test de classement)

Il y a quelques critères à respecter pour réaliser le test :

- Mode de prise de l'échantillon
 - o Il est souvent possible de comparer un nombre plus élevé d'échantillons par flairage (olfaction) qu'avec mise en bouche
- Le nombre d'échantillons doit être réduit pour des composés agressifs (ex : moutarde de Dijon)

Pour analyser ces tests, la méthode conseillée est la méthode de Friedman (méthode des rangs).

Exemple : 5 yaourts (k=5) ont été classés par 10 sujets (n=10) selon l'intensité de la saveur vanillée par ordre décroissant (1 étant le plus vanillé, le 5 le moins vanillé) (voir le tableau ci-dessous)

Yaourts	Classement					Sommes des rangs
	1	2	3	4	5	
148	6	4	0	0	0	6+4x2 = 14
920	3	3	3	1	0	22
640	0	0	0	3	7	47
825	0	1	2	4	3	39
174	1	2	5	2	0	28

On constate que le yaourt 148 a été placé 6 fois en 1^{ère} position (le plus vanillé), 4 fois en 2^{ème} position, ce qui laisse supposer que c'est le yaourt la plus vanillé. En revanche le yaourt 640 a eu un classement le moins vanillé, 3 fois en 4^{ème} position, 7 fois en 5^{ème} position. Mais on ne peut conclure que par une analyse statistique.

Pour analyser ce tableau, on procède en 2 étapes :

- On commence d'abord calculer la valeur de Friedman (Fr_{obs})

Si $Fr_{obs} \geq \chi^2$ pour un ddl et $\alpha \leq 0,05 \rightarrow$ il y a une différence significative globale. Cela veut dire qu'au moins un des yaourts a été trouvé significativement plus vanillé que les autres. Une 2^{ème} étape s'impose pour connaître les produits différents.

Si $Fr_{obs} \leq \chi^2$ pour un ddl et $\alpha \leq 0,05$, on conclut qu'il n'y a pas de différence entre les yaourts. L'analyse statistique s'arrête donc en 1 étape.

Calcul de la valeur de Friedman :

$$Fr = (12 (R_1^2 + R_2^2 \dots + R_k^2)) / (nk(k+1) - 3n(k+1))$$

R est la somme des rangs

n le nombre de sujet et k le nombre de produit

$$Fr_{obs} = 12(14^2 + 22^2 + 47^2 + 39^2 + 28^2) / (10 \times 5 \times 6) - 3 \times 10 \times 6 = 27,76$$

On consulte la table χ^2 (χ^2) de la page 14 :

Le ddl ici est : nombre de produits - 1 = 5 - 1 = 4

Extrait de la table χ^2 (page 14)

ddl	10% (α)	5%	2,5%	1 %	0,1 %
4	7.78	9.49	11.14	13.28	18.47

On constate que 27.76 (Fr_{obs}) est supérieur à 18.47. On conclut donc qu'il existe une différence significative globale au seuil de 0,1 % (ou 0,001) en intensité vanillé entre les yaourts. Si Fr_{obs} était inférieur à 9.49, on aurait conclu qu'il n'y a pas de différence entre les yaourts. L'analyse s'arrêterait à ce niveau.

La 2^{ème} étape consiste en la mise en évidence des yaourts plus vanillés que les autres. On fait appel à une comparaison multiple des différences des sommes de rangs des produits ($|R_i - R_j|$) pour déterminer des différences significatives entre les couples d'échantillons et établir un classement hiérarchique.

Si $|R_i - R_j| \geq \delta$ théorique (pour un α donné), on conclut que les produits ont été trouvés différents.

Si $|R_i - R_j| \leq \delta$ théorique (pour un α donné), on conclut que les produits ont été trouvés identiques.

Calcul de δ théorique au seuil de $\alpha=0,05$ $\delta = 1,960 \frac{\sqrt{nk(k+1)}}{6}$

Calcul de δ théorique au seuil de $\alpha=0,01$ $\delta = 2,576 \frac{\sqrt{nk(k+1)}}{6}$

Dans notre exemple, calculons δ théorique au seuil de $\alpha=0,05$ $\delta = 1,960 \frac{\sqrt{10 \cdot 5(6)}}{6} = 13,86$

Si la différence des sommes de rangs d'un yaourt par rapport à l'autre, soit $|R_i - R_j|$ est supérieure à cette valeur théorique (13,86), on conclut que ces 2 yaourts ont été trouvés différents (ici en note vanillée) avec un risque α de 0,05. Dans le cas contraire ils sont identiques en note vanillé.

On va ordonner les yaourts dans l'ordre décroissant de la somme des rangs pour faciliter les comparaisons (remarque : sans faire le test statistique, on peut supposer que plus la somme des rangs est faible, plus le yaourt est vanillé, mais cela ne suffit pas. Il faut un test statistique)

Yaourts	148 ^a	920 ^{a,b}	174 ^{b,c}	825 ^{c,d}	640 ^d
Somme des rangs (R)	14	22	28	39	47

On compare les yaourts l'un par rapport à l'autre au niveau de somme des rangs, soit $|R_i - R_j|$.

Si $|R_i - R_j| \geq 13,86$, on conclut que les yaourts ont été significativement trouvés différents ($\alpha=0,05$). Dans ce cas ces yaourts ne porteront pas la même lettre (en exposant).

Si $|R_i - R_j| \leq 13,86$, on conclut que les yaourts ont été trouvés identiques ($\alpha=0,05$). Dans ce cas ces yaourts porteront la même lettre (en exposant).

Exemples : comparons les yaourts 148 et 920 $\rightarrow 22 - 14 = 8$ $8 < 13,86$. Ces 2 échantillons ne sont pas différents, porteront donc la même lettre : a (voir le tableau)

Yaourts 148 et 174 $\rightarrow 28 - 14 = 14$; $14 > 13,86$. Ces 2 échantillons sont différents, ne porteront pas la même lettre : b (voir le tableau)

Yaourts 920 et 825 $\rightarrow 39 - 22 = 17$ $17 > 13,86$. Ces 2 échantillons sont différents, ne porteront pas la même lettre : c (voir le tableau)

Yaourts 174 et 640 $\rightarrow 47 - 28 = 19$ $19 > 13,86$. Ces 2 échantillons sont différents, ne porteront pas la même lettre : d (voir le tableau)

Pour résumer, les échantillons qui porteront la même lettre ne seront pas différents en note vanillée, contrairement aux échantillons qui portent des lettres différentes.

Conclusion :

148 > 174, 825 et 640 \rightarrow 148 étant plus vanillé que les 3 autres indiqués

640 < 174, 148 et 920 \rightarrow 640 moins vanillé que les 3 autres indiqués

148 et 920 ne sont pas différenciés.....

Ces 3 rochers ont été évalués avec 2 répétitions par 6 sujets entraînés sur le descripteur « intensité chocolat ». Chaque sujet reçoit trois rochers codés sur des assiettes. L'ordre de présentation des produits est aléatoire. Les produits sont notés sur une échelle de 10 cm de faible à forte (Faible=0, Forte=10).

Le tableau page 18 du polycopié montre les notes attribuées par les sujets, chaque cellule comporte deux valeurs qui correspondent aux répétitions des sujets. On y trouve aussi la moyenne de notes par sujet et par produit.

Sur une échelle de 10 cm, en observant la moyenne des produits on peut supposer le classement suivant en intensité chocolat : 55B (7,34)>55C(6,98)>50C(5,42)

Or pour pouvoir conclure les différences significatives en intensité chocolat entre les 3 rochers, nous devons réaliser une analyse de variance (en anglais Analysis of variance ou Anova) (Le tableau page 19 du polycopié).

Ce tableau permet de connaître les sources de variance dus aux produits, aux sujets (donc variance à 2 facteurs) ainsi qu'à l'interaction produits*sujets. (Remarque : analyse de variance faite avec un logiciel conduit à ce type de tableau)

Voir le tableau page 19 (risque α de 5%)

Si $F_{\text{calculé}} \geq F_{\text{théorique}}$ \rightarrow on conclut qu'il y a une différence significative pour le facteur considéré (produits ou sujets ou interaction produits*sujets)

Si $F_{\text{calculé}} \leq F_{\text{théorique}}$ \rightarrow on conclut qu'il n'y a pas de différence significative pour le facteur considéré (produits ou sujets ou interaction produits*sujets)

\rightarrow Comment sont-elles déterminées les valeurs de $F_{\text{théorique}}$ qui figurent dans le tableau page 19 ?:

Voir la table de la page 21 pour le risque alpha de 5 % (à noter que la table de la page 22 est utilisée pour un risque alpha de 1%).

$F_{\text{théorique}}$ pour les produits : n_1^b correspond au ddl du facteur (2 pour les produits car 3 produits-1=2) ; n_1^c correspond au ddl « interaction produits*sujets » pour le facteur produit (ici 10) . L'intersection ddl 2 et ddl 10 donne la valeur de 4.10 que l'on retrouve sur la page 19.

$F_{\text{théorique}}$ pour les sujets : n_1^b correspond au ddl du facteur (5 pour les sujets car 6 sujets-1=5) ; n_1^c correspond au ddl « variations résiduelles » pour le facteur sujet (ici 18) . L'intersection ddl 5 et ddl 18 donne la valeur de 2.77 que l'on retrouve sur la page 19.

*$F_{\text{théorique}}$ pour « interaction produits*sujets »:* n_1^b correspond au ddl du facteur (10 ici car 2*5=10) ; n_1^c correspond au ddl « variations résiduelles » pour le facteur « interaction produits*sujets » (ici 18) . L'intersection ddl 10 et ddl 18 donne la valeur de 2.41 que l'on retrouve sur la page 19.

Revenons au tableau p. 19 :

- $F_{\text{calculé}} \geq F_{\text{théorique}} (22,18 > 4,10) \rightarrow$ les produits sont différents au niveau de l'intensité de chocolat avec un risque alpha de 5%. Cela veut dire que sur les 3 produits, il y en a au moins un qui est différent des autres.
 - **Si Si** $F_{\text{calculé}} \leq F_{\text{théorique}} \rightarrow$ les produits ne sont pas significativement différents pour un risque alpha de 5%. Il existe certainement une différence, mais le risque est supérieur à 5%. Dans ce cas on conclut que les produits ne sont pas différents car le risque le plus élevé accepté est de 5 %. L'analyse statistique s'arrête.
 - La variabilité entre les sujets (2^e ligne) : $F_{\text{calculé}} \geq F_{\text{théorique}} (5,43 > 2,77) \rightarrow$ les sujets n'ont pas noté de la même façon les produits (différence significative). Cette variabilité entre les sujets est **acceptable**, à condition que les sujets aient été entraînés sur le descripteur en question. Dans le cas contraire, il faut recommencer le test avec les sujets entraînés sur le descripteur, car la signification du descripteur doit être identique pour tous les sujets.
 - 3^e ligne : interaction produit-sujet : $F_{\text{calculé}} < F_{\text{théorique}} \rightarrow$ il n'y a pas eu d'interactions entre le produit et le sujet. Ce résultat est **acceptable** à condition que les sujets aient été entraînés sur le descripteur en question. Dans le cas contraire, il faut recommencer le test avec les sujets entraînés sur le descripteur, car la signification du descripteur doit être identique pour tous les sujets.
 - Maintenant qu'on sait qu'il y a une différence significative entre les produits, on va chercher à savoir si les 3 sont différents ou si un seul. On va calculer une valeur de delta ou la ppds (plus petite différence significative). On va la comparer avec les moyennes des produits (table page 18)
- Si $i - j \geq \text{ppds}$ pour un risque alpha donné \rightarrow les produits ont été trouvés significativement différents pour le descripteur intensité chocolat (i et j = moyennes des produits).
 - Si l'inverse, les produits ont été notés identiques, pas de différence d'intensité chocolat.
 - Calcul de ppds = $t \times \sqrt{\frac{2 \times \text{CME}}{n \times r}} = 2,228 \times \sqrt{\frac{2 \times 0,56}{6 \times 2}} = 0,68$
 - t = t de student (table de la page 23)
 - Pour déterminer la valeur de t, on utilise ddl de « interaction produit-sujet » qui est 10 (table de la page 19). La table nous donne 2,228
 - CME = carré moyen des écarts de « interaction produit-sujet » qui est 0,56 (table de la page 19).
 - n = nombre de sujets (6)
 - r = nombre de répétitions (2)
 - comparer les différences des moyennes des produits versus la ppds :
 - $55B - 55C = 7,34 - 6,98 = 0,36 < \text{ppds} (0,68)$
 - Pas de différences significatives à 5% entre 55B et 55C. Ces 2 produits porteront la même lettre en exposant soit a.

- $55B - 50C = 7,34 - 5,42 = 1,92 > \text{ppds} (0,68)$
 - différences significatives à 5% entre 55B et 50 C. Ces 2 produits ne porteront pas la même lettre en exposant.
- $55C - 50C = 6,98 - 5,42 = 1,56 > \text{ppds}(0,68)$
 - différences significatives à 5% entre 55C et 50 C. Ces 2 produits ne porteront pas la même lettre en exposant.

Les résultats peuvent être présentés de façon suivante :

55B	55C	50C
7,34 ^a	6,89 ^a	5,42 ^b

Rappelons que la même lettre indique qu'il n'y a pas de différence.

Conclusion

- ➔ Les rochers 55B et 55C n'ont pas été significativement trouvés différents en intensité chocolat (risque alpha de 5%). Ils ont été trouvés plus intenses en descripteur chocolat que le rocher 50C.
- ➔ Le chocolatier peut décider de produire un rocher avec 55% de B (cacao de brésil) ou 55% de C (cacao de Côte d'ivoires) s'il souhaite se fixer uniquement sur ce descripteur.

Les pages 24-27 donnent les résultats d'une évaluation sur une échelle de 10 cm de 4 chocolats sur le descripteur « dureté » réalisés par 16 sujets (analyse par logiciel Fizz).

Le 2^{ème} tableau de la page 24 montre que les produits ont été trouvés significativement différents (risque alpha <0,0001). De même il existe une variance dû aux sujets (juges) (risque 0,0103).

Page 25 : le logiciel a calculé la ppds : 1,46 . Le tableau donne les chocolats identiques ou différents sur le descripteur « dureté ». Rappelons que la même lettre n'indique pas de différence. On peut conclure que les chocolats Lindt et Côte d'or présentent significativement (5%) plus de dureté que les chocolats Nestlé et Carrefour.

Page 26 : La comparaison des données sur le descripteur « dureté » (moyenne des notes) peut être aussi réalisée par un test Newman-Keuls (une variante de l'analyse de variance permettant de comparer moyenne vs moyenne). Ce test donne les mêmes résultats qu'avec la ppds (page 25).

Page 27 : la ppds à 5% permet de comparer les sujets en fonction de leurs réponses.

- Par ailleurs, il est possible de faire une représentation type graphique radar à partir de données de notation sur une échelle (pages 35 et 36 du photocopié)

IV.7. Epreuves descriptives

Permettent de décrire avec un minimum de mots et un maximum d'efficacité le produit à analyser, de manière à établir une carte d'identité précise, reproductible et compréhensible.

Comment est réalisée cette analyse ? Plusieurs étapes :

- Rechercher un minimum de descripteurs qui permettront de donner le maximum d'informations sur les propriétés sensorielles du produit.
- Mesurer l'intensité de la sensation perçue pour chaque descripteur choisi.
- Construire à l'aide de l'ensemble des descripteurs quantifiés le profil sensoriel du produit.

Le choix des descripteurs :

- Terminologie utilisée pour décrire doit être simple, aisément compréhensible par les sujets
- Existe déjà des listes préétablies de descripteurs pour certains produits (café, chocolat, lait, vin...)
- Ou établissement des descripteurs : ceci comporte plusieurs étapes
 - 1^{ère} étape : Recherche du plus grand nombre possible de descripteurs et sélection libre
 - On réunit le panel (au moins une 20aine de personnes), on présente les produits aux sujets et travaillent individuellement puis communément, discussion sur les descripteurs générés par chaque sujet
 - 2^e étape : Premier tri qualitatif
 - Tous les termes hédoniques (harmonieux, agréables, écœurants...) et les termes non-pertinents (rien à voir avec le produit) seront éliminés
 - 3^e étape : Deuxième tri quantitatif
 - Chaque sujet reçoit la liste des descripteurs ayant subi le premier tri. Il re-déguste les mêmes produits et attribue, à chaque descripteur, une note sur une échelle d'intensité.
 - On somme les intensités (cumulées) et on prend les descripteurs aux intensités les plus élevées.
 - 4^e étape : Troisième tri statistique
 - A partir de chaque descripteur retenu, on réalise une
 - analyse factorielle des correspondances (AFC)
 - classification ascendante hiérarchique ou dendrogramme (CAH)afin de réduire la liste.

Après ces 4 étapes, on ne retient souvent que 10% de descripteurs de la 1^{ère} étape.

Une fois qu'on a obtenu ces descripteurs :

- 5^e étape : Entraînement à la liste de descripteurs
 - Pour que chaque descripteur soit monodimensionnel et être compris par les sujets
 - Il est utile d'élaborer un lexique établissant chacun des termes descriptifs employés. Lorsque c'est possible, il faut fournir au sujet une référence concrète représentant le descripteur (exemple : pour les notes

aromatiques des solutions de molécules aromatiques : banane, vanille, fraise.....)

- 6^e étape : suivi du panel
 - En analyse sensorielle, on cherche des réponses répétables, justes et reproductibles.
 - Il est possible d'évaluer l'efficacité et performance des sujets : **coefficient de Spearman** = connaître l'aptitude d'un sujet à classer correctement des produits pour un critère donné : classement des solutions de concentrations connues (ex : solutions sucrée, salée, amère...)

Exemple pour le coefficient de Spearman (Sp) : classement d'une solution sucrée de 5 à 40 g/L dans l'ordre de sucrosité croissante (de 1 à 6)

Solution sucrée g/L	Rang attendu	Classement par sujet i	di (Rang attendu - classement sujet i)	di ²
5	1	2	1	1
10	2	3	1	1
15	3	1	2	4
20	4	4	0	0
30	5	5	0	0
40	6	6	0	0
$\sum di^2$				6

$$Sp = 1 - 6 \sum di^2 / n(n^2 - 1) \quad n = \text{nombre de produits ou solutions}$$

Pour le sujet i (voir le tableau ci-dessus)

$$Sp = 1 - (6 \cdot 6) / (6(6^2 - 1)) = 0,83$$

Si le sujet avait classé correctement toutes les solutions, Sp aurait été 1.

En faisant ce type de test, on peut constituer un groupe avec les sujets présentant des Sp élevés. Imaginons qu'un fabricant de produit alimentaire souhaiterait diminuer la teneur en sucre de son produit. On peut soumettre ses nouveaux produits en vue d'une évaluation sensorielle à un groupe performant pour détecter le goût sucré.

- Lorsque l'on dispose plusieurs variables et plusieurs données, on effectue une analyse en composante principale (ACP) (en anglais Principal component analysis=PCA) qui permet d'observer sur un graphique la disparité des sujets au niveau de l'évaluation sensorielle.

Avant de donner un exemple à ce niveau, quelques notions sur l'ACP :

Il s'agit d'une analyse statistique multi-variée et réalisée avec un logiciel. L'ACP permet d'expliquer la portion la plus grande de la variance totale. Elle permet de mettre en évidence des similarités ou oppositions entre variables et de repérer les variables les plus corrélées entre elles. Le graphique de l'ACP est souvent présenté avec 2 axes (Axe 1 ou Axe 2 ou F1 ou

F2) qui présentent les % les plus élevés de l'information (exemple : graphique de la page 29 : F1 représente 79,80 % de l'information, F2 19% de l'information. Ces % sont calculés par le logiciel, voir tableau page 28 ACP valeurs propres). Plus une variable est projetée vers l'extrémité de l'axe, plus elle pèse au niveau de l'information. Deux variables bien représentées et proches l'une de l'autre sont corrélées positivement tandis que deux variables qui s'opposent sont corrélées négativement.

Pour introduire l'ACP, prenons un exemple sur les températures mensuelles moyennes de 15 villes françaises sur 30 ans : 15 observations (villes), 12 variables (mois de l'année) (tableau de la page 28). On remarque un nombre important d'observations et de variables. L'ACP va permettre de montrer sur un graphique les informations principales (figure de la page 29).

Remarque : l'axe F1 est l'axe horizontale ; l'axe F2 l'axe verticale.

Sur le graphique de la page 29, on note les variables corrélés positivement entre elles : janvier-décembre ; février-novembre ;mai-juin-juillet-août.

Le graphique de la page 30 montre le positionnement des villes (observations ou individus) sur les 2 axes F1 et F2 : On note que les villes de Montpellier, Marseille forment un groupe, quant 'à Paris, Vichy, Clérmont un autre... Brest est bien différencié des autres villes....

Le graphique de la page 31 est la superposition des 2 graphiques précédents (observations et variables).

Un classement ascendant hiérarchique ou arbre hiérarchique (page 32) permet de regrouper les villes en fonction des températures observées sur 30 ans.

- Après cette introduction sur l'ACP, revenons au cas de l'analyse sensorielle. On avait évoqué plus haut que l'on pouvait visualiser par graphique ACP la disparité des sujets (ou juges) au niveau de l'évaluation sensorielle.
- Exemple (ACP juges, page 34) : 3 variétés de melon (Cézanne, Sirio, Cyrano) dégustés 2 fois par 12 juges sur 8 descripteurs (variables) : intensité arôme melon, intensité odeur melon, sucré, juteux, fermeté...Les chiffres sur le graphique correspondent aux numéros de juges (■). On note que les juges 24 et 28 ont tendance à noter plus fortement les descripteurs juteux, sucré, persistance fibres contrairement au sujet 34. Le juge 33 note plus fortement le descripteur fondant contrairement aux juges 23 et 65. Le juge 34 note plus fortement l'intensité couleur.

Les 2 graphiques ACP de la page 33 sur l'analyse sensorielle de 3 melons :

Le graphique ACP à gauche montre les variables (descripteurs). Les variables, intensités arôme global, odeur global, couleur ainsi que sucré et fondant sont positivement corrélées. Ces variables sont corrélées négativement à la variable ferme. Cela veut dire que les produits qui ont été notés fortement pour le variable « ferme », ont été faiblement notés pour les variables intensités arôme, odeur.....

Le graphique ACP à droite de la page 33 montre les produits (individus ou observations). On remarque que la variété Cyrano s'oppose à la variété Sirio. Si on revient sur l'ACP à gauche de cette page, on peut se renseigner sur les variables spécifiques de ces 2

variétés : la Cyrano se distingue par des descripteurs intensités fondant, intensité odeur, juteux, couleur alors que la variété Sirio se distingue par fermeté.

IV.8. Epreuve hédonique

Utilisées lorsque l'on veut recueillir l'acceptabilité et la préférence des produits auprès des consommateurs dit sujets 'naïfs'. Ces épreuves sont conduites avec un nombre important de sujets (consommateurs), qui d'après AFNOR, doit être de l'ordre de 60 (mais souvent supérieur). Ces tests peuvent se faire en entreprise, mais de préférence sur des populations ciblées.

Les tests statistiques les plus utilisés sont :

- Le test de classement (Friedman) : classer les produits (croissant/décroissant en fonction d'un critère).
 - o Il fournit une information relative sur la préférence ou l'acceptabilité des produits
- 2° test : La notation sur une échelle (pour connaître le niveau de satisfaction)
 - o Echelles de préférence structurées (de préférence)
- Tests hédoniques : caractère agréable, acceptables des produits....