

# Statistique Générale

Sébastien Lê

Les « Stat » à l'agro c'est...

## ...des projets

- *Hédonisme et catégorisation, application sur de données de vins*
- *Approches holistiques appliquées à des stimuli sonores: une étude sur des hauts parleurs de téléphones portables*

# ...des projets

- *Intimate Projective Mapping ou comment combiner contexte et émotions pour comprendre le consommateur, application à des rouges à lèvres*
- *La catégorisation libre, une nouvelle méthode de recueil de données*

## ...des projets

- *Modélisation de la température de la surface de la mer*
- *La malnutrition dans la région de Dakar*

...des prix



# Trophées Syntec des Etudes

- Vous êtes étudiant de master en marketing et études de marché, **Participez à la nouvelle édition des Trophées des Etudes Marketing & Opinion.**
- Quatre catégories de prix sont proposées:
  - Mémoire de master recherche M2: **1500 euros**
  - Mémoire de master professionnel M2: **1500 euros**
  - Mémoire de master M1 (collectif ou individuel):**1500 euros**
  - Mémoire collectif: **2500 euros**

# Résultats de la 1ère édition des Trophées Syntec des Etudes

- **Mémoire Master Collectif :**
- L'influence de la marque sur la perception du consommateur?
- Gaël Gauthier, Guilaine Perron, Cécile Perreau, Gaël Chevet



# Résultats de la 2ème édition des Trophées Syntec des Etudes

- **Mémoire Master Collectif :**
- **Les consommateurs ont la parole ! Application à douze parfums de luxe**
- Mélanie Cousin, Maelle Penven, Mathilde Philippe, Marie Toularhouat

# Résultats de la 3ème édition des Trophées Syntec des Etudes

- **Mémoire Master Collectif :**
- Quelle est l'influence de l'apport d'information telle que la liste des ingrédients et l'appellation du produit sur la perception du consommateur?
- Lucie DECRETON, Carole JEGOU, Laure LE PAIH, Magali MONNIER

# Résultats de la 4ème édition des Trophées Syntec des Etudes

- **Mémoire Master Collectif :**
- "Une nouvelle méthode de tri : le tri hiérarchique"
- SCHMIED Camille, de la PORTBARRÉ Charles, MELHINGER Stéphanie, KEMARREC Cathy, HERSCHTAL Mélusine

# Résultats de la 5ème édition des Trophées Syntec des Etudes

- Catégorie Master 1. **Affichage de l'éco-conception : approche quantitative d'un problème complexe.** Perrine Baudry et Coline Rondeau .
- Catégorie master 2, mémoire collectif. **Intimate projective mapping : comment le contexte et les états émotionnels peuvent aider à comprendre le comportement du consommateur.** Prix spécial du jury décerné à Maud Leportier, Lisa Defeyter et Charlène Benniza.

...des stages



- Recherche d'une méthode de caractérisation sensorielle du haricot vert surgelé de la marque Bonduelle.



- Participation au développement de nouveaux produits ou à l'optimisation de produits existants dans la gamme des fourrages/pâte à glacer/fourrage aqueux.



- Analyse et comparaison des résultats de tests consommateurs afin de recommander des méthodes permettant des résultats « stables ».



# Dior

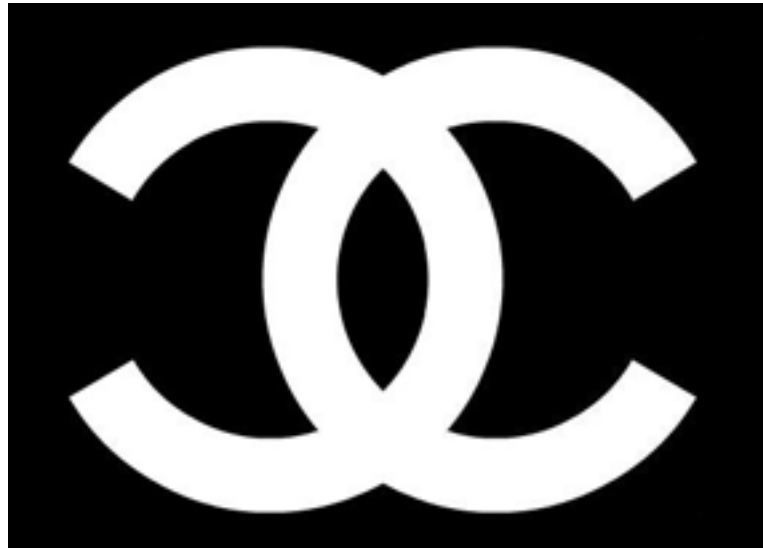
- Accompagner la mise en place d'un "Preference mapping" sur l'hydratation visant à mieux comprendre la perception de l'hydratation, sujet de fond pour les équipes de développement R&D.

# L'ORÉAL

- Analyse de la fiabilité des données recueillies et réalisation de mapping de produits sur la base de données d'évaluation.

# International Flavors & Fragrances

- Optimisation d'une approche sensorielle dans la mesure de la performance des parfums et des matières premières de parfumerie par un panel interne.



- Recherche d'une cartographie de produits (soins et maquillage).

# Inserm



**Institut national  
de la santé et de la recherche médicale**

- Identification statistique de médicaments associés aux risques d'accident de la circulation.



- Analyse des déterminants de l'exposition combinée de la pollution à des mélanges de pesticides.



- Analyse statistique de données d'écologie microbienne sous R.
- Etude statistique de la réponse de la végétation en communauté d'alpage à différents patrons spatiaux de défoliation expérimentation.

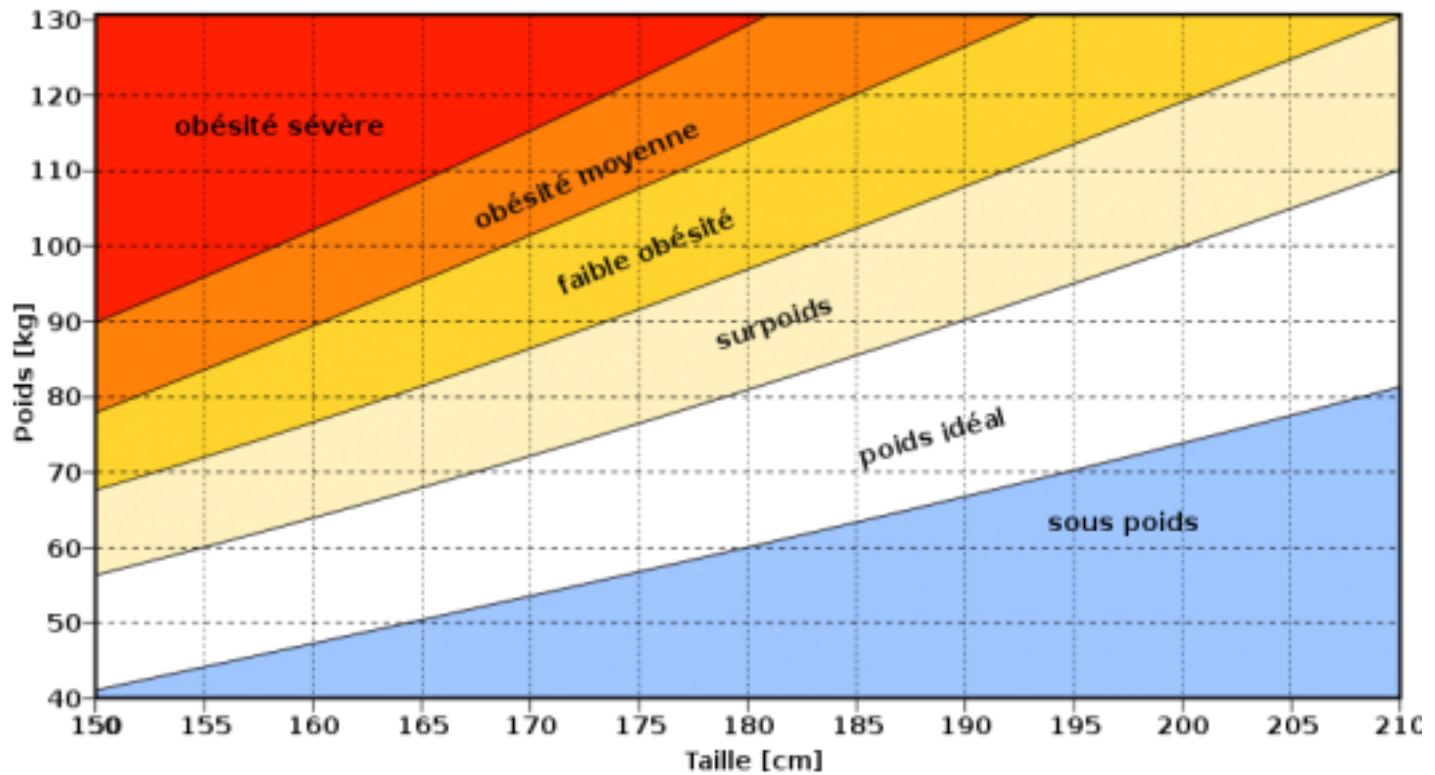
# PETIT PROBLÈME DE STATISTIQUE...



<http://bigbrowser.blog.lemonde.fr/2012/06/25/super-size-me-plus-dune-britannique-sur-deux-est-en-surpoids/>

- *"Un problème inquiétant."* C'est ainsi que le Dr Davina Deniszczyc définit à la BBC le résultat de l'étude qu'elle a menée avec d'autres chercheurs pour l'ONG caritative Nuffield Health.
- D'après les données portant sur 30 000 femmes, elle conclut que 52,5 % des femmes britanniques sont en surpoids, c'est-à-dire avec un indice de masse corporelle (IMC, obtenu en divisant le poids en kilogrammes par la taille en mètres au carré) supérieur au "poids idéal" (IMC entre 18,5 et 25) ; 16,2 % d'entre elles sont considérées obèses (IMC supérieur à 30).

<http://bigbrowser.blog.lemonde.fr/2012/06/25/super-size-me-plus-dune-britannique-sur-deux-est-en-surpoids/>



# Tour de taille moyen de 84,9 cm, 85,5 cm pour les Françaises

- L'étude constate également que 57 % des femmes britanniques ont un tour de taille plus important qu'il ne devrait pour garantir une bonne santé. Le tour de taille moyen est de 84,9 cm, quand il devrait normalement être d'environ 80 cm. C'est dans le nord de l'Angleterre qu'il est le plus important, avec 87 cm en moyenne, contre 81,9 cm à Londres.
- Davina Deniszczyc met en garde contre les conséquences de la situation : "Cela peut ressembler à un sujet cosmétique, il n'est pas question des femmes rentrant dans les jeans serrés. La graisse concentrée autour de la taille peut conduire à d'importants problèmes de santé, comme des cancers du sein et l'infertilité."
- Les femmes britanniques ne sont pas les seules concernées : en France, selon [l'enquête ObEpi de Roche avec l'institut public Inserm, publiée en 2009](#), 26 % des Françaises (et 38,5 % des Français) sont considérées en surpoids, 15,1 % étant obèses (13,9 % chez les hommes). Concernant le tour de taille moyen, celui-ci est de 85,5 cm chez les Françaises, soit 5,7 cm de plus qu'en 1997 (79,8 cm). Celui des hommes est, quant à lui, passé de 91,3 cm à 94,8 cm.

# Comment ont-ils fait?

- Sachant que la « population française » est estimée à 65 350 181 habitants (1<sup>er</sup> Janvier 2012)
- Sachant que la « population française » est composée de 51.4% de femmes
- Sachant que chaque minute 1.010464 personne décède en France
- Sachant que chaque minute 1.577915 enfant naît en France

# Démarche statistique

- Question
- Recueil
- Traitement
- Décision
- Retour éventuel à la question

# Nature des variables mesurées

- Quantitatif
- Qualitatif

# Nature du problème

- Statistique exploratoire
- Statistique inférentielle
- Population
- Echantillon

# Nature du problème

- Statistique exploratoire
- Statistique inférentielle
- Population
- Echantillon
- **Démonstration avec R**



# Suite à notre discussion...

- La question à laquelle je cherche à répondre m'amène à me situer soit dans un contexte **exploratoire**, soit dans un contexte **inférentiel**
- On ne peut pas prendre de décision sur la base d'une seule **moyenne**, il nous faut l'information sur la **variance** (ou celle sur l'écart type)

# Suite à notre discussion...

- La distribution d'échantillonnage nous permet de comprendre les propriétés de nos estimateurs
- On distingue
  - estimation ponctuelle
  - estimation par intervalle

## Suite à notre discussion...

La **moyenne**  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$

est un estimateur « naturel » de

la **moyenne**  $\mu$  (que je ne connaîtrai jamais)

$\bar{X}_n$  a en plus de bonnes propriétés

# Distribution d'échantillonnage

$$E(\bar{X}_n) = \mu$$

$$V(\bar{X}_n) = \frac{\sigma^2}{n}$$

# Estimation par intervalle de confiance (1)

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0; 1)$$

$$\left[ \bar{X}_n - u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}; \bar{X}_n + u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right]$$

# Rappels

La **variance**  $S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$

est un estimateur « naturel » de  
la **variance**  $\sigma^2$  (que je ne connaîtrai jamais)

On lui préférera  $S'^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$

qui a de bonnes propriétés

## Estimation par intervalle de confiance (2)

$$\frac{\bar{X}_n - \mu}{S'/\sqrt{n}} \sim T(n-1)$$

$$\left[ \bar{X}_n - t_{1-\alpha/2}^{(n-1)} \frac{S'}{\sqrt{n}}; \bar{X}_n + t_{1-\alpha/2}^{(n-1)} \frac{S'}{\sqrt{n}} \right]$$

# Estimation par intervalle de confiance (3)

$$V(\bar{X}_n) = \frac{\sigma^2}{n} \frac{N - n}{N - 1}$$

$$\left[ \bar{X}_n - t_{1-\alpha/2}^{(n-1)} \frac{S'}{\sqrt{n}} \sqrt{\frac{N - n}{N - 1}}; \bar{X}_n + t_{1-\alpha/2}^{(n-1)} \frac{S'}{\sqrt{n}} \sqrt{\frac{N - n}{N - 1}} \right]$$



## Suite à notre discussion...

La **variance**  $S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$

est un estimateur « naturel » de  
la **variance**  $\sigma^2$  (que je ne connaîtrai jamais)

On lui préférera  $S'^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$

qui a de bonnes propriétés

# Estimation par intervalle de confiance

$$\frac{(n-1)S'^2}{\sigma^2} \sim \chi^2(n-1)$$

$$\left[ \frac{(n-1)S'^2}{q_{1-\alpha/2}^{(n-1)}}; \frac{(n-1)S'^2}{q_{\alpha/2}^{(n-1)}} \right]$$

# Fin de la 1<sup>ère</sup> semaine

Cours 1

TD1 et TD2