

L'enseignement de la statistique dans les cursus universitaires d'économie et de gestion

Evelyne LAURENT

Centre d'Analyse et de Recherches en Economie (CARE-EA 2260 EMR)

Faculté de Droit, de Sciences Economiques et de Gestion

Université de Rouen – 76186 ROUEN CEDEX 1

evelyne.laurent@univ-rouen.fr

Trois objectifs principaux:

1. permettre la description des phénomènes économiques tant au niveau macroéconomique qu'au niveau des agents économiques.
2. éclairer le choix du décideur confronté à l'incertitude.
3. établir des relations entre des variables qui décrivent et mesurent des phénomènes et des comportements.

1 - Tout d'abord une finalité descriptive, tant au niveau macroéconomique qu'au niveau des entreprises et des ménages.

Exemples pour les entreprises :

- suivi des quantités produites.
- présentation de l'évolution des ventes par secteurs, par produits.
- étude de la pyramide des âges, de la répartition des salaires des employés de l'entreprise.
- etc.

Exemples des principaux agrégats macroéconomiques concernant :

- l'investissement, de la production des entreprises.
- la consommation des ménages.
- les échanges extérieurs.
- le chômage.
- etc.

2 - Les méthodes statistiques permettent d'éclairer le choix du décideur confronté à l'incertitude.

2.1 - D'une part, les applications des lois de probabilité sont nombreuses dans l'entreprise.

Exemples:

- déterminer un stock de sécurité pour un produit en modélisant sa demande.
- gérer les files d'attente en modélisant le nombre de clients se présentant à un guichet.
- prévoir le taux de pénétration d'un nouveau produit.

Les lois exponentielle, logistique, normale et de Poisson sont les plus fréquemment utilisées.

2.2 - D'autre part, la théorie de l'estimation et les tests paramétriques trouvent un maximum d'applications en économie de l'entreprise.

- dans les études de marché précédant par exemple le lancement d'un produit nouveau ;
- dans les choix d'investissement ;
- dans l'estimation et le contrôle sur échantillon de la qualité d'une fabrication.
- etc.

3 - Le troisième objectif est d'établir des relations entre des variables qui décrivent et aussi mesurent des phénomènes et des comportements.

L'outil de référence est **l'économétrie** qui apparaît :

- comme un **outil de validation** (ou d'invalidation) de théories économiques, commerciales, etc...
- comme un **outil d'investigation** :
 - ✓ en mesurant les conséquences de la variation de la valeur d'une variable sur une ou plusieurs autres variables : ex : quel impact l'augmentation du revenu national d'un pays va-t-elle avoir sur la consommation des ménages ? (**simulation**)
 - ✓ en anticipant les réactions des agents économiques face à des mesures de politique économique : quelles seraient les répercussions d'une baisse du taux d'imposition en matière de consommation des ménages et/ou d'investissement des entreprises ? (**prévision**)

Les différentes étapes dans l'apprentissage:

- 1) statistique descriptive
- 2) calcul élémentaire des probabilités
et lois statistiques
- 3) statistique inférentielle
- 4) économétrie.

Difficultés récurrentes dans :

- l'utilisation
- la compréhension
- la maîtrise des outils proposés

La statistique descriptive :

une matière facile ?

- Bien souvent les étudiants se contentent d'un résultat numérique sans avoir la moindre idée sur sa signification.

La connaissance d'une formule et son application correcte leur semblent amplement suffisantes...

- En outre, la confiance absolue qu'ils ont en leur calculatrice (et en la manière dont ils s'en servent) les détourne de la réflexion sur la cohérence du résultat obtenu.

Exemples de points de butée :

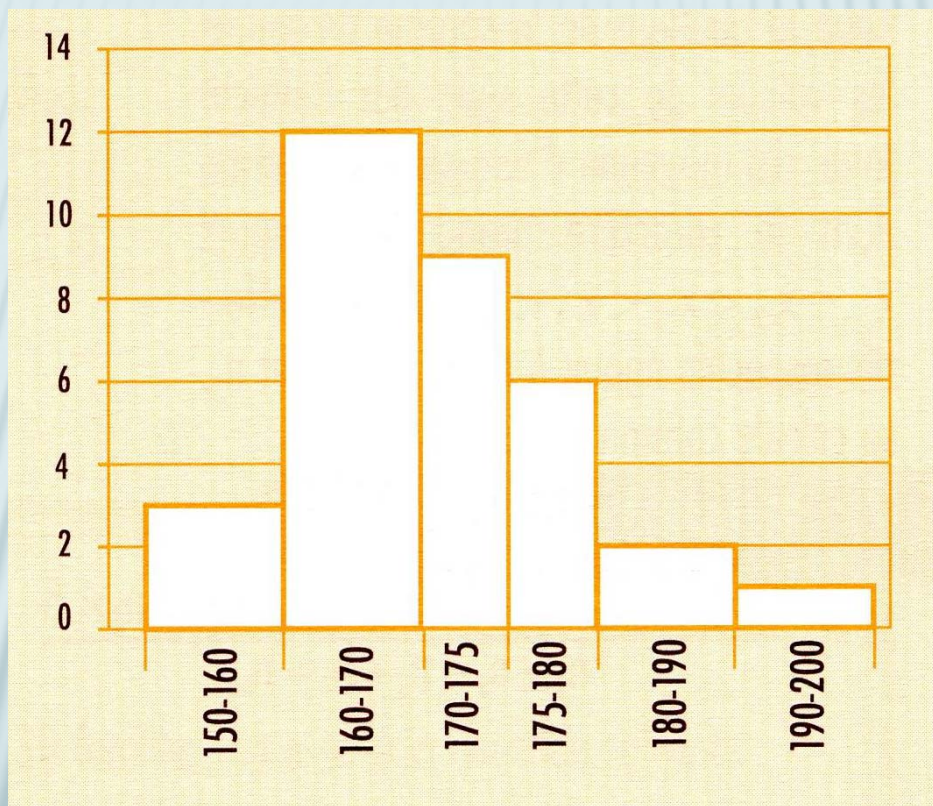
- La pondération : pas de problème lorsqu'on présente la moyenne arithmétique mais semble plus difficile à saisir pour d'autres paramètres (autres moyennes, variance, etc.)
- Les représentations graphiques: l'histogramme. Intégrer le fait que l'aire d'un rectangle est proportionnelle à l'effectif (ou à la fréquence d'une classe).

Histogramme

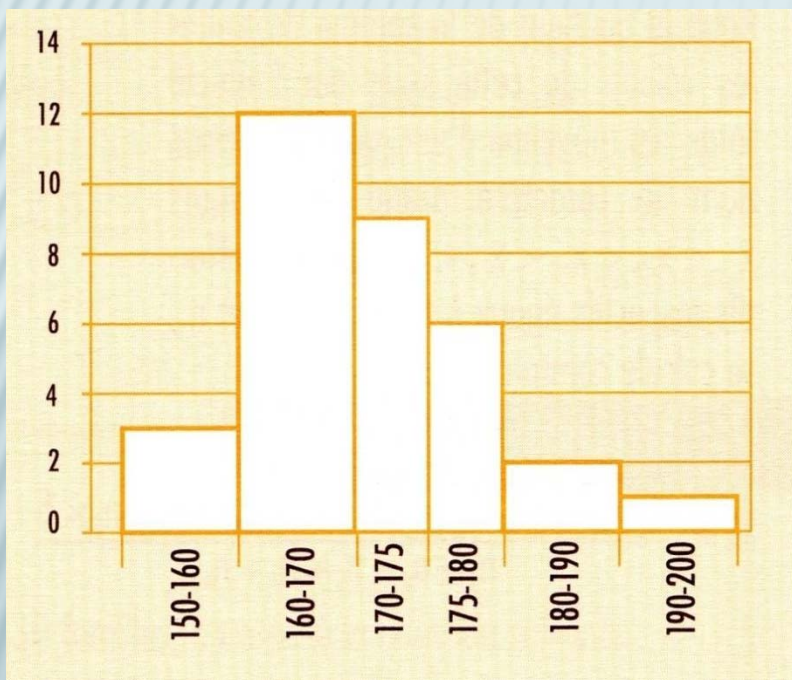
Vous devez savoir que si le plus souvent une série statistique est donnée par valeur, comme dans l'exemple précédent, il peut aussi arriver qu'elle soit donnée par « classe de valeur ». Ainsi, on peut consigner dans un tableau comme suit les classes de tailles des individus d'un groupe et les représenter sous la forme d'un histogramme :

Extrait de: *Je révise les maths, cahier de vacances de 10 à 110 ans*,
Editions CNRS, 2009

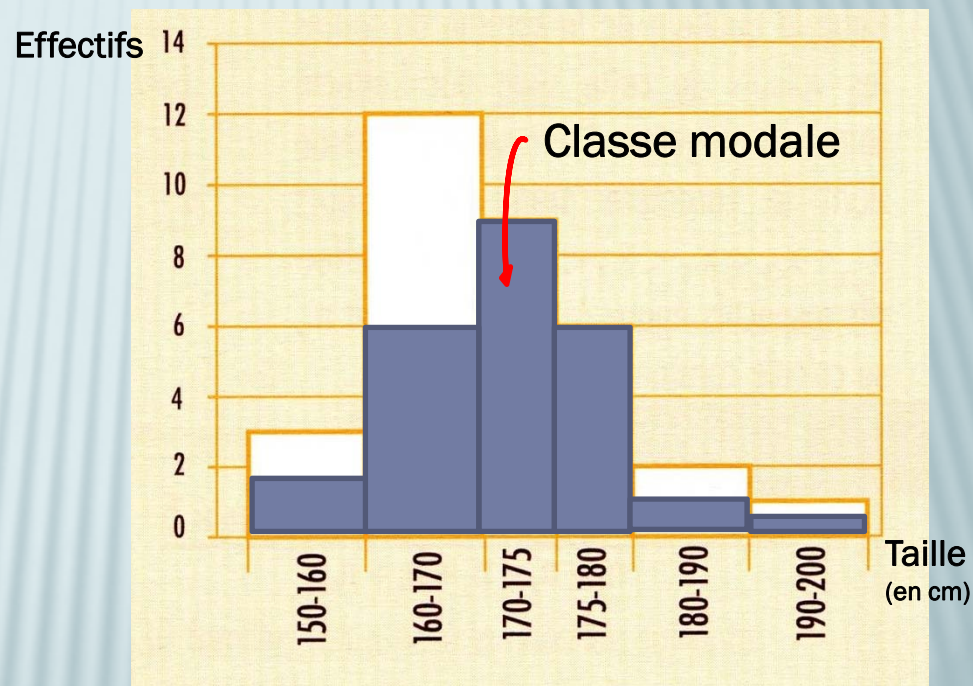
Tailles	150-160	160-170	170-175	175-180	180-190	190-200
Effectifs	3	12	9	6	2	1



Tailles	150-160	160-170	170-175	175-180	180-190	190-200
Effectifs	3	12	9	6	2	1



FAUX



CORRECT

Statistique et calcul des probabilités: une liaison dangereuse...

**Le passage de la statistique descriptive
(domaine du certain)**

**à la statistique mathématique
(domaine de l'incertain),**

avec une formulation mathématique plus poussée,

déroute bon nombre d'étudiants.

Etapes de l'enseignement:

- Axiomes fondamentaux du calcul des probabilités
- Variables aléatoires
- Lois statistiques

Comme en statistique descriptive, les formules sont apprises par cœur, mais la difficulté réside dans l'utilisation judicieuse de la boîte à outils.

Un minimum de réflexion est nécessaire, surtout sur des exercices de synthèse.

Il est habituel d'aborder la notion de probabilité à partir de la fréquence.

Mais cette approche fréquentiste induit parfois des raisonnements erronés comme celui-ci:

Alex se présente à un concours où, sur 800 candidats, il y a 40 admis :

« La probabilité que je réussisse ce concours n'est que de 5%... »

Non Alex, les admis au concours ne sont pas tirés au hasard...

On peut relever un certain nombre d'erreurs classiques par rapport à ce programme, reflétant une mauvaise maîtrise de certains notions et entraînant des confusions comme par exemple :

- *Entre $(A \cap B)$ et $(A \cup B)$*
- *Entre A et B indépendants et A et B incompatibles.*
- *Entre $P(A \cap B)$ et $P_B(A)$*
- *Etc...*

La notion d'indépendance (d'événements, de variables aléatoires), est souvent mal appréhendée.

Exemple:

On étudie une population selon 2 caractères: le poids et la taille. La ventilation des individus est faite selon 3 classes de taille T_1 , T_2 et T_3 et 2 classes de poids, P_1 et P_2 .

On donne: $P(T_1) = 20\%$, $P(T_2) = 50\%$, $P(T_3) = 30\%$.

A l'intérieur de chaque classe de taille, les probabilités pour un individu d'appartenir à la classe de poids P_1 sont les suivantes:

$P_{T_1}(P_1) = 60\%$, $P_{T_2}(P_1) = 45\%$, $P_{T_3}(P_1) = 35\%$.

- 1) Calculer les probabilités de chacun des couples possibles.*
- 2) Calculer les probabilités de chaque poids.*
- 3) En déduire les probabilités d'appartenance à chaque classe de taille à l'intérieur de chaque classe de poids.*

Autre difficulté récurrente:

$$\textit{Entre } Y = nX \quad \textit{et} \quad Y = \sum_{i=1}^n X_i$$

Exemple : Soient 10 personnes montant dans un ascenseur.

Le poids d'une personne est une variable aléatoire X avec $E(X) = 75$ kg et $\sigma_x = 10$ kg.

Soit Y la variable aléatoire égale au poids total des 10 personnes.

Calculer l'espérance et l'écart-type de Y .

Avant de calculer ces paramètres, il convient d'écrire la relation (correcte) entre Y et X .

L'expérience nous a montré qu'il faut s'attendre à trouver la relation $Y = 10 X$

alors qu'il convient d'écrire :

$$Y = \sum_{i=1}^{10} X_i$$

X_i : poids de la personne i .

Difficultés propres aux lois statistiques:

- l'indépendance des épreuves, permettant l'utilisation de la loi binomiale, est souvent mal maîtrisée ou utilisée à mauvais escient.
- l'utilisation de la table de la fonction de répartition de la loi normale.

La statistique inférentielle :

autres difficultés !

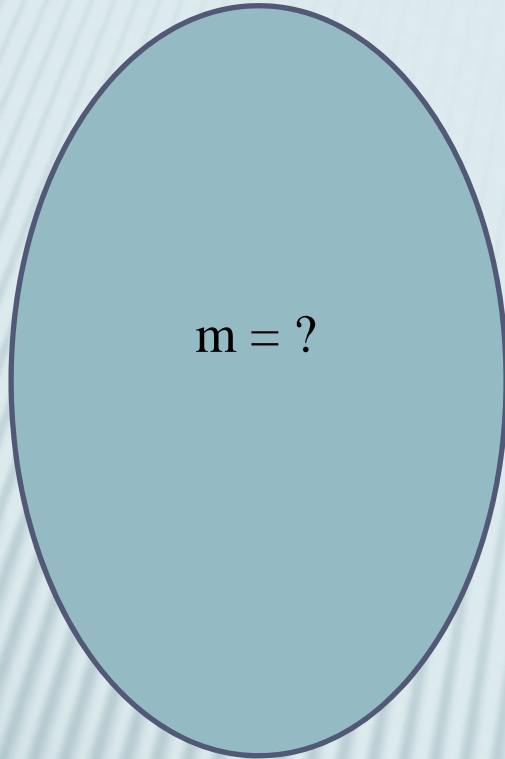
Une difficulté globale d'assimilation de la démarche :

- la distinction entre d'une part la population-mère, les valeurs prises par certains paramètres sur celle-ci, et d'autre part l'échantillon et les valeurs des paramètres sur celui-ci semble une notion purement théorique dont la portée est loin d'être saisie.
- ainsi une erreur fréquemment commise consiste à confondre valeur réelle et valeur estimée d'un paramètre.

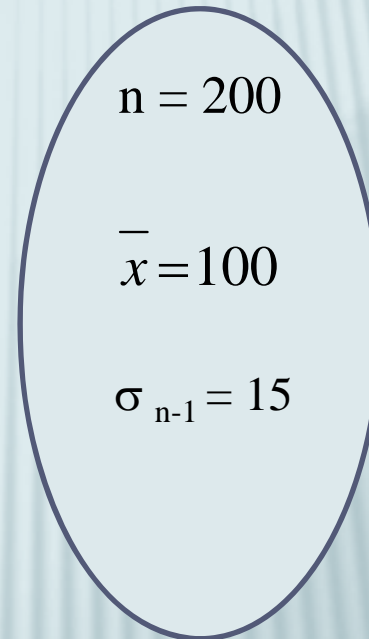
Exemple-type de confusion relatif à la théorie de l'estimation:

Au vu des observations effectuées sur un échantillon, il est demandé de donner une estimation pour la moyenne arithmétique (inconnue !) de la population-mère.

Population-mère



Echantillon



L'intervalle de confiance est généralement construit correctement mais souvent avec une conclusion pour le moins surprenante :

« l'échantillon est bien représentatif de la population-mère » !

Une autre difficulté réside dans la formulation des tests.

Si le test n'est pas explicitement posé dans l'énoncé de l'exercice,

ex : tester $H_0 : p = p_0$ contre $H_1 : p > p_0$

les étudiants sont rapidement déroutés.

Difficultés également dans le rapprochement entre les concepts économiques (développés en cours d'économie) et les techniques statistiques.

Par exemple, l'estimation des paramètres a et b sur une fonction de consommation keynésienne:

Consommation des ménages = a [Revenu des ménages] + b

D'autre part, la difficulté de l'enseignement de la statistique réside également dans le degré de théorisation qu'il convient d'adopter pour un public de non-spécialistes.

Si la formalisation mathématique apparaît indispensable dans la présentation et la justification des concepts, certaines démonstrations un peu ardues paraissent devoir être écartées.

En conclusion, nous pouvons noter:

- **l'importance de la statistique dans les cursus d'économie et de gestion, tant les applications sont nombreuses.**
- **la nécessité pour les enseignants chargés des enseignements de la statistique de faire preuve de beaucoup de pédagogie.**

Merci de votre attention...