

Analyse de données et modélisation statistique
(Pour le gestionnaire)

Franck Moraux.

franck.moraux@univ-rennes1.fr

Cours : Franck MORAUX.

TD d'analyse de données : Dany VIT.

TD de modélisation : Emmanuelle FROMONT.

A. OBJECTIF

Ce cours est consacré aux outils les plus courants d'analyse et de modélisation statistiques utilisés aujourd'hui par les gestionnaires. Il aborde différentes méthodes statistiques sous un angle théorique, tout en accompagnant l'étudiant dans la résolution de problèmes très opérationnels et pratiques.

B. ORGANISATION

Ce cours s'organise autour de 6 séances de cours magistraux (de 4 heures) et de 6 TD de 2 heures chacun. De nombreux exemples, exercices et applications numériques seront traités tant en cours qu'en TD. Les TD se dérouleront essentiellement en salle informatique. Notons qu'on a aussi prévu en cours quelques démonstrations en « temps réel », de nombreux exemples devant illustrer la théorie exposée de façon traditionnelle. L'utilisation de tableaux doit permettre à l'étudiant(e) de mettre en pratique les notions théoriques vues en cours. Au terme du cours, l'étudiant(e) devra pouvoir a) formuler un modèle mathématique pour chacun des modèles; b) effectuer un traitement adéquat des données ; c) interpréter les résultats des analyses effectuées et le cas échéant valider les hypothèses de modélisation.

C. CONTENU

1. **Généralités et rappel sur les méthodes statistiques.** Typologie des variables. Relations entre variables. Notion de corrélation et mesures. Coefficients de détermination partielle. Démarche de modélisation.

Les **modèles de régression** les plus usuels en gestion seront d'abord abordés :

2. **Régression simple sur une variable exogène.** Principe. Méthode des Moindres Carrés Ordinaires. Propriétés des estimateurs. Inférence (estimations ponctuelle et par intervalle) et tests sur les coefficients. Analyse de la variance et qualité de la régression. **Analyse des résidus.** Test de normalité. Hétéroscédasticité (effets, détection et tests). Autocorrélation (effets, détection et tests).
3. **Régression multiple sur plusieurs variables exogènes.** Estimation et tests. Qualité et Choix du modèle. Méthodes de sélection (ascendante, descendante, par étapes). La colinéarité et ses effets. La régression classique sur des variables indicatrices ou *dummies*. La prévision et la modélisation des séries temporelles. **Approfondissement¹** : Maximum

¹ pour information.

de Vraisemblance et tests de Wald, LM et LR. Régression de variable qualitative sur des variables quantitatives.

L'**analyse de données** sera représentée par 3 approches :

4. **L'analyse en Composante Principale.** Les données traitées et la matrice de corrélation. Test de sphéricité de Barlett. « Test » de KMO (Kaiser-Meyer-Olkin), Le choix du nombre d'axes à retenir. L'interprétation des axes factoriels. L'interprétation des contributions. Les coordonnées des individus – les scores factoriels. Les rotations.
5. **L'analyse Factorielle des Correspondances.** Les données traitées et les tableaux de fréquence. Les tableaux de profils (associés aux) lignes, de profils (associés aux) colonnes. Le principe de l'AFC. Traitement du tableau profils-lignes. La métrique du Chi-deux et l'inertie. Représentations graphiques. L'analyse des correspondances multiples : ACM. Le principe. Les nuages de points
6. **La typologie.** Les données analysées. Le principe. La mesure de proximités entre individus (Cas des variables quantitatives, des variables qualitatives. Les méthodes de classification hiérarchiques (CAH). Les méthodes de classification ascendantes non hiérarchiques (CANH). L'interprétation des résultats). Visualisation des groupes obtenus. Validation de la typologie. La quantité globale d'informations restituée par la typologie. Qualité relative de la typologie

D. BIBLIOGRAPHIE (à titre indicatif)

Ouvrages généralistes :

- Tenenhaus M. : *Méthodes statistiques en gestion*. 2nde édition, Dunod, 2004.
- Saporta : *Probabilités, analyse de données et statistique*. Technip Ed., 1990.

Ouvrages sur la modélisation statistique.

Manuel de base :

- Bourbonnais R. : *Econométrie*. 3^{ème} édition, Dunod, 2000.

Manuel d'approfondissement :

- Johnston J., Dinardo J. : *Méthodes économétriques*. 4^{ème} édition, Economica, 1999.

Ouvrages sur l'analyse de données :

Manuel de base :

- Evrard, Pras, Roux : *Market : Etudes et recherche en marketing*. Dunod, 2001.