

# PLAN DE COURS

## STT2400 - Régression linéaire - 3 crédits (Hiver 2012)

**Professeur(e):** Pierre Lafaye de Micheaux  
*Coordonnées:* Bureau: 4249  
Courriel: [lafaye@dms.umontreal.ca](mailto:lafaye@dms.umontreal.ca)  
Téléphone: 514-343-6607  
Web: <http://www.biostatisticien.eu/moodle>  
*Disponibilités:* Vendredi 11:30-13:30

**Assistant(es) d'enseignement:** Folly Adjogou  
*Coordonnées:* Bureau(x): 6156  
Courriel(s): [adjogou@dms.umontreal.ca](mailto:adjogou@dms.umontreal.ca)  
*Disponibilités:* Mardi 10:30-12:30

### 1 Description succincte du cours

Méthode des moindres carrés. Théorèmes de Gauss-Markov et de Cochran. Estimation et tests d'hypothèses. Résidus et diagnostics. Construction de modèles. Exemples. Utilisation du progiciel SAS.

### 2 Horaire et format du cours

Cours théoriques : 39h  
Travaux pratiques : 24h

Activité	Jour	De - À	Du - Au	Nb séances	Local	Immeuble
Cours	Vendredi	10:30-11:30 14:30-16:30	6 janvier - 2 mars	9+9	Z-240	Pav. C-McNicoll
			16 mars - 30 mars	3+3		
			13 avril	1+1		
TP	Mardi	08:30-10:30	10 janvier - 14 février	6	Z-240	Pav. C-McNicoll
			28 février	1		
			13 mars - 10 avril	5		

Chaque séance de cours prendra la forme d'un exposé magistral assuré par le professeur. Ces séances seront parsemées de périodes de résolution d'exercices. Les séances de travaux pratiques (TP) seront assurées par l'assistant d'enseignement. Ces séances permettront à l'étudiant de mettre en pratique, à l'aide de l'outil informatique (logiciels SAS et R), et également par la résolution d'exercices, les notions théoriques vues en cours.

### 3 Approche pédagogique et accès au site web du cours

La première méthode pédagogique repose sur l'exposé magistral (exposé professoral hebdomadaire). Accompagnés de périodes de résolution d'exercices en classe et de résolution d'exercices à faire à la maison, les exposés magistraux feront acquérir des connaissances théoriques et empiriques et un bagage conceptuel qui constitueront le coeur de notre enseignement. La deuxième méthode est la mise en pratique des apprentissages

par les étudiants, d'une part lors des séances de travaux pratiques, et d'autre part dans la réalisation des devoirs maison et aux examens. Cette mise en pratique impliquera plusieurs niveaux cognitifs visés par notre enseignement : compréhension, application des connaissances acquises, analyse et synthèse, sens critique.

L'étudiant pourra également utiliser certains outils mis à disposition sur le site Internet *moodle* associé à ce cours : forum de discussion, notes de cours, etc. Ainsi, l'étudiant qui désire poser des questions sur certains points du cours, des exercices ou des TPs qu'il n'aurait pas bien compris, devra soit les poser pendant les périodes de disponibilités, soit utiliser le forum *moodle* associé à ce cours. Le professeur et le démonstrateur **ne répondront à aucune question de ce type par courriel ou par téléphone.**

Notez enfin que pour accéder au site *moodle* de ce cours, **vous devez être inscrit** au cours. L'accès au site se fait via l'adresse [www.biostatisticien.eu/moodle](http://www.biostatisticien.eu/moodle) et nécessite d'utiliser votre identifiant/mot de passe :  
Identifiant : votre code d'accès DGTIC (ex : p0123456)  
Mot de passe : changeme (à changer à la première connexion)

**Remarque :** un certain délai (1 semaine) peut être nécessaire entre le moment où vous êtes inscrit au registrariat et où vous avez accès au site *moodle*. Si vous n'êtes pas capable de vous connecter **après ce délai**, merci de m'envoyer un courriel contenant les informations suivantes : Nom, Prénom, Code d'accès DGTIC, Code permanent (3 premières lettre du nom ...), Sigle du cours (STT2400), adresse courriel **obligatoirement** de l'UdeM (...@umontreal.ca).

## 4 Évaluation des apprentissages

Deux devoirs : 20% (remettre le 1<sup>er</sup> devoir avant le 14 février 8:45 et le 2<sup>ième</sup> devoir avant le 10 avril 8:45).  
Examen intra : 35% (le mardi 21 février de 8:30 à 10:30 au 2010, Pav. M. Youville).  
Examen final : 45% (*probablement* le mardi 24 avril de 9:30 à 12:30, au 1140, Pav. A.-Aisenstadt).

### Le détail des épreuves :

- *Devoirs maison :* Les devoirs peuvent être réalisés en équipe de **deux personnes maximum**. Tout **retard non justifié**<sup>1</sup> dans la remise d'un devoir entraînera la non correction de celui-ci et la note 0 sera alors donnée pour ce devoir. L'objectif du devoir maison est double. D'une part, il permet de se préparer en vue des examens (intra et final), d'autre part il permet de s'assurer que l'on maîtrise bien les éléments théoriques et pratiques vus en cours et en TP. Par ailleurs, il s'agit également d'une occasion pour l'étudiant de développer sa maîtrise de la rédaction d'un texte. Ainsi, il sera demandé à l'étudiant de soigner la présentation du travail rendu (en impression **recto-verso**, si possible sur papier recyclé) et la qualité de l'argumentation. L'étudiant **motivé** et possédant certaines aptitudes en informatique et/ou souhaitant faire une maîtrise plus tard, pourrait envisager de compléter son devoir à l'aide du logiciel libre et gratuit de traitement de texte évolué L<sup>A</sup>T<sub>E</sub>X. Une documentation sur ce logiciel est disponible sur le site internet du professeur, dans la rubrique «Manuels en ligne». Le *code source* des énoncés des devoirs maison sera rendu disponible sur le site internet du cours et l'étudiant pourra y ajouter sa contribution (les réponses aux questions) dans les parties réservées à cet effet (signalées explicitement dans le code source). Un bonus pourra être attribué à l'étudiant faisant preuve d'audace dans ce domaine!
- *Examens en classe :* Les examens *intra* et *final* se dérouleront en classe et porteront sur les exposés professoraux, les devoirs maison, les séances de travaux pratiques et les exercices du livre support. Les étudiants auront droit à une page aide mémoire (1 page au format letter 8.5x11 pouces, recto-verso où ne figure aucun exercice ou exemple résolu) qui sera donnée par le professeur la semaine précédant l'examen et qui sera également distribuée le jour de l'examen. Il est indispensable que l'étudiant amène sa calculatrice durant les examens.

---

<sup>1</sup>Sur la définition du motif valable de retard d'absence à une évaluation ; consulter le guide pédagogique de la FAS. <http://www.etudes.umontreal.ca/reglements/fas.html>

Veillez noter que la date limite pour abandonner le cours «sans frais» est le **19 janvier**. Vous pouvez abandonner le cours «avec frais» (sans être remboursé) mais sans échec inscrit au bulletin jusqu'au **11 mars**. Par la suite, si vous abandonnez, vous aurez un échec indiqué sur votre bulletin de notes. Vous avez également l'obligation de motiver une absence prévisible à l'intra ou au final dès que vous serez en mesure de constater que vous ne pourrez être présent. Il appartiendra à l'autorité compétente de déterminer si le motif est acceptable.

### Critères généraux d'évaluation :

- Contenu : acquisition des connaissances et de leur application adéquate, capacité d'analyse et de synthèse, cohérence de l'argumentaire, capacité de relier les sujets entre eux, capacité d'évaluer, de faire preuve d'un sens critique appuyé par les connaissances ;
- Forme : respect des directives afférentes aux diverses épreuves (travail, examen), respect des règles de présentation matérielle (page de couverture, prescriptions de longueur, etc...) ;
- Qualité rédactionnelle : orthographe, grammaire, ponctuation, clarté dans l'expression ;
- Ponctualité : le respect des délais prescrits.

### Remarques importantes :

- La présence en classe est **fortement recommandée** pour bien réussir ce cours !
- La note des devoirs contribuera à la moyenne finale uniquement si l'étudiant(e) obtient au moins 50% pour la moyenne pondérée des deux examens.
- Tout plagiat<sup>2</sup> entraînera de graves sanctions.
- Seules les notes des deux devoirs maison et de l'intra seront affichées sur le site du cours.

## 5 Objectifs

À la fin du cours, l'étudiant devrait :

- ✓ maîtriser les concepts clés de la méthode de régression linéaire simple et multiple ;
- ✓ avoir développé un regard critique sur cet outil ;
- ✓ interpréter une analyse produite par SAS ou R ;
- ✓ avoir acquis des connaissances informatiques suffisantes pour être en mesure de pouvoir effectuer l'analyse d'un jeu de données à l'aide des routines SAS ou R pour la régression linéaire ;
- ✓ être en mesure de produire un rapport décrivant les résultats obtenus.

## 6 Plan du cours

### 6.0 Rappels et notations

Types de variables. Résumés numériques d'une variable quantitative. Résumés graphiques d'une variable quantitative. Coefficient de corrélation. Nuage de points. Population versus échantillon. Modèle. Simulation. Notations : variable aléatoire, déterministe, etc. Estimateurs (ponctuels et par intervalle) d'un paramètre inconnu. Critères de qualité d'un estimateur. Tests d'hypothèse.

### 6.1 La régression linéaire simple

Introduction. Modélisation mathématique. Modélisation statistique. Estimateurs des moindres carrés. Interprétations géométriques. Inférence statistique. Exemples.

### 6.2 La régression linéaire multiple

Introduction. Modélisation. Estimateurs des moindres carrés. Interprétation géométrique. Exemples.

---

<sup>2</sup>Voir <http://www.fas.umontreal.ca/plagiat>

### 6.3 Inférence dans le modèle gaussien

Estimateurs du maximum de vraisemblance. Nouvelles propriétés statistiques. Intervalles et régions de confiance. Exemple. Prédiction. Les tests d'hypothèse. Exemples.

### 6.4 Validation du modèle

Analyse des résidus, de la matrice de projection. Effet d'une observation, variable explicative.

### 6.5 Régression sur variables qualitatives

Introduction. Analyse de la covariance. Liens avec l'ANOVA.

### 6.6 Choix de variables

Introduction. Colinéarité (VIF). Choix incorrect de variables : conséquences. La sélection de variables en pratique. Critères classiques de choix de modèles. Procédure de sélection. Exemple.

## 7 Livre support obligatoire

Cornillon, P.-A. et Matzner-Lober, E. (2010). *Régression avec R*, Springer.

## 8 Livres fortement recommandés

Ces ouvrages sont disponibles à la bibliothèque Maths-Info : <http://www.bib.umontreal.ca/MI>

- Lafaye de Micheaux, P., Drouilhet, R. et Liquet, B. (2010). *Le logiciel R - Maîtriser le langage - Effectuer des analyses statistiques*, Springer.  
<http://www.springerlink.com/content/978-2-8178-0114-8>
- Cornillon, P.-A. et Matzner-Lober, E. (2007). *Régression. Théorie et applications*, Springer.  
<http://www.springerlink.com/content/978-2-287-39692-2>
- Lejeune, M. (2010). *Statistique - La théorie et ses applications, 2<sup>ème</sup> édition*, Springer.  
<http://www.springerlink.com/content/978-2-8178-0156-8>

*Dernière mise à jour: 4 janvier 2012.*