

PLAN DE COURS

STT1700 - Introduction à la statistique - 3 crédits (Automne 2012)

Professeur(e): Pierre Lafaye de Micheaux

Coordonnées: Bureau: 4249
Courriel: lafaye@dms.umontreal.ca
Téléphone: 514-343-6607
Web: <http://www.biostatisticien.eu>

Disponibilités: Jeudi 13:00-15:00

Assistant(es) d'enseignement: M.-O. Billette ; P. Bodson-Clermont ;
A. Deschênes ; G. Elgbeili ;
J. Sango ; A.-A. Vallée

Coordonnées: 4257 ; 4160 ;
Bureau(x): 4160 ; 5246 ;
6146 ; 5264 .
marc-olivier.billette@umontreal.ca ; pm.bodson-clermont@umontreal.ca ;
Courriel(s): alexandre.deschenes@umontreal.ca; guillaume.elgbeili@umontreal.ca ;
joel.sango@umontreal.ca; audrey-anne.vallee@umontreal.ca
mercredi 12h20-14h30 (AA-4257) ; mardi 9h-10h, jeudi 12h45-13h45 (AA-4160)

Disponibilités: ??? (AA-???) ; jeudi 13h30-15h30 (AA-5246)
mardi 15h30-17h30 (AA-6146) ; mardi 8h-10h (AA-5264)

1 Description succincte du cours

Description des données. Production de données. Probabilités. Inférence. Intervalles de confiance et tests d'hypothèses. Données de dénombrement. Tableaux de contingence. Régression linéaire simple.

Remarques : Utilisation du logiciel libre et gratuit R.

2 Horaire et format du cours

Cours théoriques : 36h

Travaux pratiques : 16h/Gr.

Chaque séance de cours prendra la forme d'un exposé magistral assuré par le professeur. Ces séances seront parsemées de périodes courtes de résolution de problèmes. 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 d'exercices, les notions théoriques vues en cours.

Le tableau qui suit présente les horaires du cours.¹

¹Pour les horaires officiels, consulter le site <http://www.etudes.umontreal.ca/horaire>

Activité	Jour	De - À	Du → Au	Nb séances	Local	Immeuble
Cours	Jeudi	08:30-10:30	4 octobre → 20 décembre	12	B-2245	Pav. 3200 J-B
Cours	Vendredi	11:30-12:30	5 octobre → 2 novembre	5	B-2245	Pav. 3200 J-B
			9 novembre 16 novembre → 21 décembre	1 6	S1-151 B-2245	Pav. Jean-Coutu Pav. 3200 J-B
TP Gr.1	Lundi	13:30-15:30	15 octobre → 22 octobre	2	B-4285	Pav. 3200 J-B
			5 novembre	1	B-4285	
			12 novembre	1	B-3255	
			19 novembre	1	B-4285	
			3 décembre → 17 décembre	3	B-4285	
TP Gr.2	Lundi	13:30-15:30	15 octobre → 22 octobre	2	B-4290	Pav. 3200 J-B
			5 novembre → 19 novembre	3		
			3 décembre → 17 décembre	3		
TP Gr.3	Lundi	13:30-15:30	15 octobre → 22 octobre	2	B-3280	Pav. 3200 J-B
			5 novembre → 19 novembre	3		
			3 décembre → 17 décembre	3		
TP Gr.4	Lundi	13:30-15:30	15 octobre → 22 octobre	2	B-3275	Pav. 3200 J-B
			5 novembre → 19 novembre	3		
			3 décembre → 17 décembre	3		
TP Gr.5	Lundi	13:30-15:30	15 octobre	1	Z-330	Pav. C-McNicoll
			22 octobre	1	Z-215	Pav. C-McNicoll
			5 novembre	1	Z-300	Pav. C-McNicoll
			12 novembre	1	Z-215	Pav. C-McNicoll
			19 novembre	1	N-515	Pav. Roger-Gaudry
			3 décembre → 17 décembre	3	Z-215	Pav. C-McNicoll
TP Gr.6	Lundi	15:30-17:30	15 octobre → 22 octobre	2	Z-245	Pav. C-McNicoll
			5 novembre	1		
			12 novembre	1		
			19 novembre	1		
			3 décembre → 17 décembre	3		

Répartition par groupe :

Gr.1 : de **Aaaa*** à **Bouo*** ; Gr.2 : de **Boup*** à **Fonu*** ; Gr.3 : de **Fonv*** à **Lacb***

Gr.4 : de **Lacc*** à **Mbeo*** ; Gr.5 : de **Mbep*** à **Samt*** ; Gr.6 : de **Samu*** à **Zzzz***

3 Approche pédagogique et accès au site web du cours

La première méthode pédagogique repose sur l'exposé magistral (exposé professoral bi-hebdomadaire). Accompagnés de périodes courtes de résolution de problèmes 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 cœur de l'enseignement dispensé. 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 aux deux examens intra-session. Cette mise en pratique impliquera plusieurs niveaux cognitifs visés par notre enseignement : compréhension, application des connaissances acquises, analyse et synthèse, sens critique.

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 à la fin du cours, soit en TP. Le professeur et les démonstrateurs **ne répondront à aucune question de ce type par courriel ou par téléphone**. L'étudiant pourra consulter la planification des exercices à faire en TP sur le site internet associé au cours.

Des notes de cours seront disponibles sur *Studium*. Il peut être très utile de les télécharger, de les imprimer et de les amener en cours car je suivrai leur contenu religieusement.

4 Évaluation des apprentissages

- Examen intra-session 1 : 25% de la note du cours (aura lieu le lundi 29 octobre de 13:30 à 15:30).
- Examen intra-session 2 : 25% de la note du cours (aura lieu le lundi 26 novembre de 13:30 à 15:30).
- Examen final : 50% de la note du cours (aura lieu le lundi 14 janvier de 13:30 à 16:30).

Le détail des épreuves :

Les deux examens *intra-session* et l'examen *final* se dérouleront en classe et porteront sur les **exposés** professoraux, les séances de **travaux pratiques** et les **exercices** à faire à la maison. 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 (via le site web) 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. Il est **TRÈS IMPORTANT** de noter que la résolution des exercices à faire pour les TPs pourra vous assurer la maîtrise et la rapidité d'exécution de certains calculs simples et répétitifs, mais il s'agit d'une activité insuffisante pour vous assurer une note honorable à l'examen. **Il vous faut donc impérativement travailler les exercices supplémentaires, et avoir compris en profondeur les notes de cours et les exemples qui y sont présentés.**

Veillez noter que la **date limite** pour abandonner le cours «sans frais» est le **19 septembre**. Vous pouvez abandonner le cours «avec frais» (sans être remboursé) mais sans échec inscrit au bulletin jusqu'au **9 novembre**. Par la suite, si vous abandonnez, la mention «Échec» sera indiquée sur votre bulletin de notes. Vous avez également l'obligation de motiver une absence prévisible à l'examen intra-session ou à l'examen final dès que vous serez en mesure de constater que vous ne pourrez être présent. Il appartiendra à l'autorité compétente (à savoir le responsable des études) 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, 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 **très fortement recommandée** pour bien réussir ce cours !
- Tout plagiat² entraînera de graves sanctions.
- Seules les notes des deux examens intra-session seront affichées sur le site du cours.

5 Objectifs

Le monde qui nous entoure est rempli de phénomènes aléatoires, que ce soit la durée de vie d'une personne, le nombre de processus qui roulent en même temps sur un ordinateur à un moment donné, le nombre de bières consommées par des étudiants un vendredi soir, etc. Certaines variables peuvent avoir un effet sur ces phénomènes. Par exemple, la consommation de bières est-elle la même chez les étudiantes que chez les étudiants ?

La statistique est la science qui permet de prendre des décisions face à l'incertitude. Pour y arriver, il

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

faut recueillir des données provenant du phénomène aléatoire et analyser ces données à partir d'un modèle stochastique qui représente le phénomène. La science de la statistique étudie les façons les plus efficaces d'effectuer la cueillette, la modélisation et l'analyse des données afin d'extraire le maximum d'information de celles-ci. Ce cours offre un survol des principales questions auxquelles la statistique tente de répondre. L'accent sera mis davantage sur la compréhension des concepts statistiques que sur le développement d'un livre de recettes statistiques. Pour ceux dont ce sera le seul cours de statistique, cette introduction vous donnera une idée des principaux enjeux de la discipline et vous permettra d'être de meilleurs consommateurs de statistiques publiées dans les journaux ou les revues scientifiques. Pour les autres, notez que les notions vues dans ce premier cours seront approfondies dans plusieurs cours subséquents. Cette introduction permettra ainsi à ceux qui poursuivront en statistique de se faire une bonne idée globale de la discipline.

6 Plan du cours

6.1 Statistiques descriptives

Variables et distributions. Types d'une variable. Diagramme en tuyaux d'orgues. Diagramme circulaire. Diagramme en bâtons. Histogramme. Tiges et feuilles. Graphique temporel. *Box plot*. Mesures de tendance centrale. Mesures de dispersion. Couples de variables. Distribution conjointe, marginale et conditionnelle. Indépendance. Nuage de points. Covariance, corrélation et droite des moindres carrés.

6.2 Probabilités et outils biostatistiques

Définitions. Axiomes et propriétés. Opérations ensemblistes. Événements impossibles, certains, disjoints. Probabilité conditionnelle. Théorème de Bayes. Événements indépendants. Vrais positifs, faux positifs, faux négatifs, vrais négatifs. Sensibilité et spécificité. Valeurs prédictives positive et négative, risque relatif, prévalence. Espace ROC.

6.3 Variables aléatoires

Définition. Variables aléatoires discrètes. Fonction de masse. Fonction de répartition. Espérance et variance d'une (somme, produit de) variable(s) aléatoire(s). Loi Binomiale. Loi multinomiale.

6.4 Lois continues

Variables aléatoires continues. Fonction de densité et fonction de répartition. Espérance et variance. Loi normale et propriétés. Théorème limite central. Approximation normale de la loi binomiale. Lois du chi-deux. Lois de Student.

6.5 Estimation ponctuelle

Population et échantillon. Échantillon aléatoire simple. Statistique et estimateur. Estimateur sans biais. Estimateur sans biais d'une moyenne, d'une variance et d'une proportion. Variance d'un estimateur.

6.6 Intervalles de confiance

Définition et interprétation. Intervalle de confiance pour une moyenne, pour la différence de deux moyennes. Intervalle de confiance pour une proportion, pour la différence de deux proportions.

6.7 Tests d'hypothèses pour de grands échantillons

Définitions. Composantes d'un test d'hypothèse : hypothèse nulle et alternative, région critique, erreurs de première et seconde espèce, valeur- p , statistique de test. Test unilatéral, bilatéral. Puissance d'un test. Tests d'hypothèses pour une moyenne, pour la différence de deux moyennes. Courbes ROC. Tests d'hypothèses pour données appariées. Tests d'hypothèses pour une proportion, pour la différence de deux proportions. Randomisation.

6.8 Inférence pour de petits échantillons

Intervalles de confiance et tests d'hypothèses pour une moyenne, pour la différence de deux moyennes.

6.9 Tests du khi-deux

Tests d'ajustement. Tests d'indépendance. Causalité et dépendance.

6.10 Régression linéaire simple

Définitions. Modèle de régression linéaire simple. Estimation des paramètres. Prédiction. Intervalles de confiance. Coefficient de corrélation.

7 Livre support obligatoire³

Mendenhall, W., Beaver, R.J., Beaver, B.M. et Ahmed, S.E. (2011). *Introduction to Probability and Statistics*, Second Canadian Edition. Nelson, Canada.

8 Livres recommandés³

- 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>
- Moore, D.S. et McCabe, J.A. (2006). *Introduction to the Practice of Statistics*, 5th edition. W. H. Freeman and Company, New York.

Dernière mise à jour: 22 octobre 2012.

³Pour la disponibilité des livres en bibliothèque, contacter le comptoir de prêt (<http://www.bib.umontreal.ca/nous-joindre/MI.htm>) ou la bibliothécaire Ferroudja Nazef (f.nazef@umontreal.ca).