

Faculté des arts et des sciences
Département de sciences biologiques

Sigle du cours	BIO 2043 B	Trimestre Automne 2014
Titre du cours	Statistique pour sciences de la vie	
Crédits	3	
Horaire	Dates : jeudi 4 septembre au jeudi 11 décembre; cours: 13h00-17h00 Salle : 2220, Pav. Liliane-de-Stewart	

Professeur	Daniel Borcard (théorie), Maryse Robert (démonstration)	
Local	F-227 PMV	
Courriel	daniel.borcard@umontreal.ca , maryse4444@hotmail.com	
Téléphone	(514) 343-6111(1233)	Télécopieur (514) 343-2293

Politique sur la durée des examens :

Un temps de battement de 15 minutes est nécessaire afin de permettre aux étudiants de déposer leur copie d'examen et de libérer la salle. Ex : examen d'une durée de 1h45 ou de 2h45

ÉVALUATION				
Examens	Pondération	Date	Durée	Salle
Intra	45%	30 octobre	2h45	2220, Pav. Liliane-de-Stewart
Final	55%	11 décembre	2h45	2010, Pav. Marguerite d'Youville

Les examens sont à choix multiples, sans ordinateur. **Documentation admise (papier seulement)**: les deux documents d'introduction au langage R (Intro_R_1.pdf et Intro_R_2.pdf) et, pour **chaque** module, **deux** feuilles de papier au format lettre, écrits ou imprimés des deux côtés. Le contenu de ces documents (résumé de la matière) est écrit par l'étudiant.

BUT DU COURS

Le but du cours est de fournir des connaissances de base en analyse statistique des données issues des sciences de la vie et de la santé (unidimensionnelles et bidimensionnelles), incluant la capacité de décider quelle méthode statistique s'applique à chaque situation particulière. L'étudiant apprendra à réaliser ses calculs à l'aide du langage statistique R. Il les expérimentera en classe pendant le cours et à l'aide de séries d'exercices formatifs.

OBJECTIFS GÉNÉRAUX D'APPRENTISSAGE

À la fin du cours, l'étudiant devra posséder les connaissances suivantes: capacité de présenter et résumer des données; capacité de choisir la ou les méthodes d'analyse appropriées à la question scientifique qui se présente, en comprendre les fondements théoriques, réaliser les calculs à l'aide du langage **R** et interpréter correctement les résultats. L'autonomie et l'appropriation des méthodes sont des buts majeurs de ce cours.

SYLLABUS GÉNÉRAL

PLAN DU COURS

Date	Modules (Sujets traités)
4 sept.	1) Introduction; définitions. Introduction au langage statistique R (1).
11 sept.	2) Échantillonnage et plans d'expérience. Introduction au langage statistique R (2).
18 sept.	3) Paramètres d'une distribution.
25 sept.	4) Lois de distribution: binomiale, normale. Lois dérivées de la loi normale: F , t , χ^2 .
2 oct.	5) Théorie de l'estimation: distribution de paramètres; intervalles de confiance (proportion, moyenne et variance).
9 oct.	6) Théorie de la décision: comment tester l'hypothèse principale? Tests paramétriques, non paramétriques et par permutation. Les deux types d'erreur; faux positifs et faux négatifs.
16 oct.	7) Comparaison des variances de deux échantillons indépendants (test F). Comparaison des moyennes de deux échantillons (tests t pour données indépendantes et appariées); tests non paramétriques correspondants.
30 oct.	Examen intra (matière des cours 1 à 6).
6 nov.	8) Corrélation de Pearson et de Kendall.
13 nov.	9) Régression linéaire simple et multiple.
20 nov.	10) Analyse de variance (ANOVA) à un critère de classification; test de Kruskal-Wallis. Test d'homogénéité des variances. Comparaisons multiples.
27 nov.	11) ANOVA à deux critères de classification de modèle I; facteurs fixes et aléatoires; ANOVA hiérarchique; ANOVA pour mesures répétées.
4 déc.	12) Variables catégorielles: les tests khi-carré.
11 déc.	Examen final (matière des cours 7 à 12).

ORDINATEUR EN CLASSE

Il est **très important** que le plus grand nombre possible d'étudiants viennent en classe avec un ordinateur portable. Celui-ci peut être un Mac (Mac OSX 10.6 ou plus récent) ou un PC (Windows XP ou 7), voire un Linux, mais pas une tablette. Des instructions sont fournies avant le début de la session pour que chacun arrive dès le premier cours avec le langage statistique R installé sur sa machine.

Le cours se donne en interaction avec les étudiants: des opérations sont réalisées en langage R afin de présenter la matière de façon active.

Pour les étudiants qui n'ont pas d'ordinateur portable personnel, il y a 10 postes de travail en accès libre au local B-342, Pav. Marie-Victorin. Ces ordinateurs ont R installé. Les étudiants peuvent donc y faire les exercices, il suffit d'y apporter une clé USB contenant les fichiers nécessaires. Dans R sur ces ordinateurs-là, il faut désigner un répertoire de travail lors de chaque nouvelle ouverture de R.

SÉRIES D'EXERCICES

Des exercices formatifs seront fournis au fur et à mesure de l'avance du cours. Ces exercices seront réalisés de manière autonome par l'étudiant, qui pourra, s'il y a lieu, poser des questions au professeur ou au démonstrateur durant la dernière heure de chaque rencontre ou par courriel. Les corrigés seront fournis à temps pour la préparation des examens, mais pas immédiatement afin d'inciter chacun à faire les exercices.

RÉFÉRENCES ET DOCUMENTATION

Manuels de référence suggérés:

Aho, Ken A. 2014. Foundational and applied statistics for biologists using R. CRC Press, Boca Raton FL, USA.

Daniels, W.W. 2009. Biostatistics. A foundation for analysis in the health sciences. 9e édition, John Wiley & Sons Ed., Hoboken NJ, USA.

Motulsky, H. 2013. Biostatistique. Une approche intuitive. Traduction de la 2e édition anglaise par M. Dramaix-Wilmet. De Boek Ed., Bruxelles, Belgique.

Samuels, M. L. & Witmer, J. A., Schaffner, A. 2010. Statistics for the life sciences, 4th Edition. Prentice Hall, Upper Saddle River NJ, USA.

Scherrer, B. 2007. Biostatistique, Volume 1. Gaëtan Morin Ed., Boucherville, QC.

Scherrer, B. 2009. Biostatistique, Volume 2. Gaëtan Morin Ed., Boucherville, QC.

Des compléments utiles en statistique avancée (multidimensionnelle) pourront être trouvés dans les livres suivants:

Borcard, D., Gillet, F. & Legendre, P. Numerical Ecology with R. UseR! Series, Springer Ed., New York, USA.

Legendre P. & L. Legendre. 2012. Numerical ecology. 3rd English ed. Elsevier Ed., Amsterdam, Pays-Bas.

Les notes de cours théoriques (PowerPoint et documents connexes), les séries d'exercices et les données correspondantes sont distribuées sur **StudiUM**.

Discussion par courriel : les étudiants peuvent écrire au professeur ou au démonstrateur par courriel pour poser des questions sur la matière. Pour poser des questions en dehors des heures de cours et de tp, on **doit** passer par cette voie, afin que tout le monde profite des échanges. Les enseignants envoient les réponses à tous les étudiants inscrits au cours.

RESSOURCES DOCUMENTAIRES DE LA BIBLIOTHÈQUE ÉPC BIOLOGIE :

Pour trouver des livres, rapports, documents audiovisuels ou localiser des livres de la réserve de cours
Catalogue Atrium : www.bib.umontreal.ca/Atrium/

Pour accéder à une base de données et chercher des articles scientifiques
Répertoire Maestro : www.bib.umontreal.ca/Maestro/ (catégorie Sciences /sous-catégorie Sciences biologiques)

Pour consulter des guides sur les ressources en bibliothèque et une sélection de sites Web en biologie
Ressources en sciences biologiques : www.bib.umontreal.ca/ED/disciplines/themabio.htm

Pour apprendre rapidement comment initier une recherche documentaire
Guide d'aide à la recherche : www.bib.umontreal.ca/ED/disciplines/biologie/guide-recherche-bio.pdf