

Faculté des arts et des sciences
Département de sciences biologiques

Sigle du cours et section	BIO2042	Hiver 2018
Titre du cours	Biostatistique II	
Crédits	3	
Horaire	Théorie : Lundi, 13h00 - 15h00 Pratique : Lundi, 15h00 - 18h00 Salle : B-4260; Pav. 3200 Jean-Brillant	

Professeur	Louis-Etienne Robert
Local	F-062 Mardi 11h00-16h00.
Courriel	louis-etienne.robert@umontreal.ca
Téléphone	514-343-6111 ext. 1304
Démos	Julian Wittische julian.wittische@umontreal.ca (Vendredi 13h00 à 15h00, Local F-062), Frédéric Dwyer-Samuel

ÉVALUATION			
Examens	Pondération	Date d'attribution	Date de remise
Intra	30 %	26 février 2018	5 mars 2018; 23h59
Rapport d'analyse (Équipe de 2)	10 %	16 avril 2018	18 avril 2018 23h59
Examen Sommatif	38 %	23 avril 2018	30 avril 2018; 23h59
Travaux pratique	22 %	-	À la fin de chaque cours-

Les examens prendront la forme de deux examens à la maison. Ces examens comporteront des questions sur l'analyse et la manipulation des données, de la programmation en R, et des critiques de lectures. Les étudiants auront une semaine pour les compléter. Les rapports seront remis à la date indiquée ci-dessus avant 23h59 sur Studium. Les remises de travaux en retard seront pénalisées à un taux de 1% par heure.

Il est attendu que les étudiants apportent leur propre ordinateur avec R (3.3.2) et Rstudio Desktop installés. Veuillez informer le professeur si vous n'avez pas accès à un ordinateur portable pour ces séances d'informatique. R et Rstudio peuvent être téléchargés **librement et sans frais** à ces adresses: <https://cran.r-project.org/> et <https://www.rstudio.com/products/rstudio/download/>

BUT DU COURS

Le but du cours est de présenter aux étudiants une portion des méthodes et techniques d'analyse statistique qui sont utilisés actuellement dans les domaines de l'écologie et de la biologie. Le cours met l'accent sur l'importance de l'exploration des données, préalables à l'ensemble des analyses statistique, ainsi que sur les modèles linéaires **généraux** (i.e., la régression linéaire et l'ANOVA) et **généralisés** (i.e., la régression logistique et de Poisson). Les étudiants acquerront également de l'expérience dans l'utilisation du langage de programmation statistique R pour entreprendre des manipulations de données et des analyses statistiques.

Chaque cours débutera avec environ deux heures de théorie et se poursuivra avec 2-3 heures de pratique. Durant ces heures de travaux pratiques, les étudiants feront sur place (et sur leur ordinateur) des exercices formatifs, en R, portant sur le sujet du cours. Une note sera attribuée à ces exercices à soumettre à la fin de la classe (2%/semaine).

OBJECTIFS GÉNÉRAUX D'APPRENTISSAGE

À la fin du cours, l'étudiant devra répondre aux objectifs suivants :

Exploratoire :

- Effectuer de l'exploration de données afin de déterminer les problèmes potentiels et les solutions statistiques possibles
- Choisir la ou les méthodes d'analyses appropriées à ses problèmes
- Lire et critiquer des articles scientifiques incluant des analyses statistiques.

Technique :


- Comprendre les fondements théoriques des modèles linéaires généraux et généralisés (LM et GLM) et appliquer ces analyses dans le contexte approprié.
- Comprendre et appliquer les outils de diagnostics ainsi que de validation des modèles linéaires généraux et généralisés
- Interpréter les résultats des modèles linéaires généraux et généralisés
- Connaître les bases de la programmation en R afin d'effectuer des manipulations de données
- S'initier aux structures de contrôles avec le langage R

SYLLABUS GÉNÉRAL

Semaine	Date	Sujet	Lecture
1	8 jan.	Introduction au cours et à la modélisation statistique. Rappel sur la corrélation et la régression linéaire. Rappel sur la programmation en R	Motulsky Ch. 32 – 33
2	15 jan.	Suppositions de la régression linéaire; diagnostics	Motulsky Ch. 34; Dalgaard Ch. 6
3	22 jan.	La régression multiple et polynomiale.	Motulsky Ch. 37, pp 296 - 305; Dalgaard Ch. 11
4	29 jan.	ANOVA avec deux critères de classification, ANOVA hiérarchique	Motulsky Ch. 39; Dalgaard Ch. 12.1 - 12.6
5	5 fev.	ANCOVA; conception expérimentale	Whitlock & Schluter Ch. 18; Dalgaard Ch. 12.7 - 12.8
6	12 fev.	Comment identifier un bon modèle ? La multicolinéarité.	Motulsky Ch. 38
7	19 fev	Comparaison des modèles avec AIC, validation croisée.	Johnson and Omlad, 2004.
-	26 fev	Synthèse/Questions 15h00 : Attribution DE L'EXAMEN INTRA	Whitman <i>et al.</i> 2004
-	5 mars	SEMAINE DE LECTURE	
8	12 mars	Modèles linéaires généralisés (GLMs); la famille exponentielle, estimation des paramètres par le maximum de vraisemblance (« <i>maximum likelihood</i> »)	Whitlock & Schluter Ch. 20
9	19 mars	La régression logistique I	Motulsky Ch. 37, pp 305 – 308 Dalgaard Ch. 13
10	26 mars	La régression logistique II, validation avec les « matrices de confusion »; ROC	Motulsky Ch. 42
11	9 avril	La régression Poisson pour les données de compte	Motulsky Ch. 6; Dalgaard Ch. 15
12	16 avril	Pour aller plus loin : Modèle non-linéaire, modèles mixtes, modèle additif généralisé, approche Bayésienne. TP : Attribution de la question/données pour le rapport d'analyse:	Motulsky Ch. 44; Dalgaard Ch. 16
-	23 avril	Synthèse/Questions 15h00 : Attribution de L'EXAMEN FINAL	

RÉFÉRENCES ET DOCUMENTATION

- Motulsky. **Biostatistique – Une Approche Intuitive**, 2me ed. De Boeck. *En réserve à la bibliothèque EPC copies des chapitres numériques sont disponible sur StudiUM.*
- Dalgaard. **Introductory Statistics with R**. *Copie numérique disponible sur StudiUM.*
- Whitlock et Schluter. 2015. **The Analysis of Biological Data**, 2nd Ed. Roberts and Company. *En réserve à la bibliothèque EPC ; copies des chapitres numériques sont disponible sur StudiUM*

 <i>Bibliothèque ÉPC-Biologie</i>
Guide en Sciences biologiques (point de départ, ressources utiles, astuces) http://guides.bib.umontreal.ca/disciplines/6--Sciences-biologiques
Recherche dans le catalogue Atrium (livres, thèses UdeM, audiovisuel, titres de revues) http://atrium.umontreal.ca/
Recherche dans les bases de données (articles scientifiques, statistiques, thèses, etc.) http://www.bib.umontreal.ca/Maestro

PLAGIAT

Nous vous invitons à consulter le règlement disciplinaire de l'Université sur le site suivant :
<http://www.fas.umontreal.ca/plagiat/>

NB : ce plan est sujet aux changements