



MAT1070M

LOGICIELS SCIENTIFIQUES - R

JEUDI 14H-16H OU 16H15-18H15 SUIVANT LE GROUPE

Enseignant A. PERRUT

Em@il Anne.Perrut@univ-lyon1.fr

Page web <http://math.univ-lyon1.fr/~perrut>

Bureau Braconnier 128 (premier étage)

Office hours Prendre rendez-vous par mail

Tél 04 26 23 45 56

Description du cours : Ce cours a pour objet une pratique avancée du logiciel R, ainsi que de l'interface R-studio. Dans le même temps, nous illustrerons par des simulations tous les aspects du cours de probabilité suivi en parallèle (principe et pratique de la simulation de variables aléatoires, calcul des probabilités, théorèmes limites, chaînes de Markov).

Prérequis : Représentations graphiques de base utilisées en statistiques, quelques notions de R.

Matériel : Les étudiant-e-s doivent installer sur leur propre ordinateur, R et R-studio, qui sont deux logiciels libres disponibles là :

<http://cran.r-project.org>

<http://www.rstudio.com>

Contexte de ce cours : il s'agit d'un cours de 12 heures, inclus dans une UE de 3 crédits proposées aux étudiant-e-s de master 1 *mathématiques appliquées, statistique* au semestre d'automne. L'autre partie du cours est consacrée à Matlab.

Objectifs du cours : À la fin de ce cours, les étudiant-e-s doivent être capable de

1. utiliser R de manière fluide pour les statistiques descriptives
2. coder des programmes plus évolués en R (boucles, création de fonctions...)
3. rédiger un rapport sous word constituant une première étude descriptive d'un tableau de données
4. connaître plusieurs méthodes de simulation de variables aléatoires, les mettre en œuvre sous R
5. illustrer la loi des grands nombres et le théorème central limite par des simulations
6. simuler une trajectoire pour une chaîne de Markov (espace d'états fini ou infini)
7. illustrer la convergence d'une chaîne de Markov vers une probabilité invariante

Évaluation (40% de l'UE) :

TP 1 : statistiques descriptives	15%
TP 2 : simulation aléatoire	15%
TP 3 : chaînes de Markov	10%
Examen final (écrit)	60%

Comportement attendu :

- L'assiduité est obligatoire dans ce cours. L'appel sera fait à chaque séance. En cas d'absence justifiée, les étudiant-e-s doivent rattraper le travail non fait par eux-mêmes.
- Les téléphones portables ne doivent pas être utilisés pendant les cours. Aucun retard en cours ne sera accepté.
- Les étudiant-e-s doivent travailler individuellement. Donner les solutions à d'autres étudiant-e-s ou accepter les solutions d'autres étudiants est un acte de plagiat qui sera puni. Les discussions entre étudiant-e-s sont encouragées, mais n'hésitez pas à poser directement vos questions aux enseignant-e-s. Par ailleurs, les étudiant-e-s doivent clairement établir ce qui relève de leur travail personnel ou ce qui relève de citations, même dans le cas de documents/programmes libres de droit. En cas de plagiat, les sanctions peuvent être variées : l'enseignant-e peut demander à l'étudiant-e de refaire le travail demandé ou de faire un autre travail, il-elle peut donner la note 0 à ce contrôle, et en cas de récurrence, l'enseignant-e peut saisir la commission de discipline.
- Les travaux (compte-rendus de TP, projets) rendus en retard ne seront pas acceptés.
- Aucun document n'est autorisé à l'examen.

Bibliographie :

Statistiques avec R, Pierre-André Cornillon et al., 2012.

Modélisation stochastique et simulation, Bernard Bercu et Djalil Chafaï, Dunod 2007.

Le logiciel R, Pierre Lafaye de Micheaux, Rémy Drouilhet et Benoît Liquet, 2014.

Planning de travail :

Ce planning est susceptible d'être modifié suivant l'avancement du groupe.

Date	Programme
Jeudi 28 sept. 2017	<ul style="list-style-type: none"> — Prise en main de R et Rstudio, statistiques descriptives (indicateurs, graphiques), coder des fonctions. — À faire : installer les deux logiciels R et R-studio. — À lire : le cours 1.
Jeudi 5 oct. 2017	<ul style="list-style-type: none"> — Les techniques de simulation des variables aléatoires (méthode du rejet, découpage d'intervalles, inverse de la fonction de répartition). Application aux lois usuelles. — À faire : finir le TP 1. — À lire : le cours 2.
Jeudi 12 oct. 2017	<ul style="list-style-type: none"> — Les techniques de simulation des variables aléatoires (II) — À faire : rendre le projet 1. — À regarder : https://www.khanacademy.org/computing/computer-science/cryptography/crypt/v/random-vs-pseudorandom-number-generators
Jeudi 19 oct. 2017	<ul style="list-style-type: none"> — Illustration des théorèmes limites par des simulations (loi des grands nombres, théorème central limite) — À lire : le cours de probabilité sur les théorèmes limites — À regarder : https://www.youtube.com/watch?v=HRnYFpDR8WM et https://www.youtube.com/watch?v=4dhm2QAA2x4
Jeudi 26 oct. 2017	<ul style="list-style-type: none"> — Théorèmes limites (II) — À lire : http://www.fil.univ-lille1.fr/~bouillaguet/HECI/cours/08_rnd.pdf
Jeudi 2 nov. 2017	<ul style="list-style-type: none"> — Vacances — Suggestion : regarder <i>Margin Call</i>, le film de J. C. Chandor.
Jeudi 9 nov. 2017	<ul style="list-style-type: none"> — TP de statistiques (estimation) — À faire : rendre le projet 2.
Jeudi 16 nov. 2017	<ul style="list-style-type: none"> — Forum Entreprises — À faire : figoler son CV et prévoir une tenue adaptée.
Jeudi 30 nov. 2017	<ul style="list-style-type: none"> — La simulation des chaînes de Markov. — À faire : relire le cours sur les chaînes de Markov
Jeudi 14 déc. 2017	<ul style="list-style-type: none"> — À faire : rendre le projet 3.
Mardi 16 jan. 2018, 13H00	<ul style="list-style-type: none"> — Examen écrit

Les jeudis 9 et 23 novembre, 7, 14 et 21 décembre et 11 janvier seront consacrés aux travaux pratiques de statistique inférentielle.