

Heure et salle

- Heure : Mercredi, 9h45—11h15
- Salle : 6.02 (sauf exception)

Enseignants :

- Joseph Salmon joseph.salmon@umontpellier.fr (Cours / TD : L2 Série C gr A3 / TP : L2 Série C gr D)
- Emmanuel Bonnet emmanuel.bonnet@umontpellier.fr (TD : L2 Série C gr D, E / TP : L2 Série C gr E)
- Tiffany Cherchi tiffany.cherchi@umontpellier.fr (TP : L2 Série C gr D)
- Pierre-Louis Montagard pierre-louis.montagard@umontpellier.fr (TD : L2 Série C gr CMI-IEN + C / TP : L2 Série C gr A3, CMI-IEN + C)
- Léa Pradier lea.pradier@etu.umontpellier.fr (TP : L2 Série C gr E)
- Raphael Romero raphael.romero@umontpellier.fr (TP : L2 Série C gr CMI-IEN + C)

Horaires de consultation (*office hours*) :

- Mercredi 17h-19h, sur rendez-vous par mail uniquement.

Prérequis et ressource pour se mettre à niveau

Les étudiants sont supposés connaître les bases de la théorie des probabilités et d'algèbre linéaire.

Probabilités : Pour le démarrage en probabilité les ouvrages suivants sont en français [CGDM01], [Ouv08], [Ouv07], [GS01].

Algèbre Linéaire : Il faut être suffisamment à l'aise avec le calcul matriciel pour bien commencer avec les modèles linéaires et aller jusqu'aux modèles économétriques : [Sch05] et [Gv13] peuvent être un bon début.

Les classiques du genre sont [Gv13, HJ94] et [TB97] est aussi un bon point d'entrée.

Algorithmique Les étudiants doivent aussi être capables d'implémenter des méthodes numériques basiques en utilisant un langage de haut niveau : Python sera la choix par défaut **MAIS** le groupe **E**, avec Léa Pradier et Emmanuel Bonnet, utilisera le langage **R**. **Tous les autres groupes** travaillerons sous **Python**.

Pour commencer en Python, un fichier d'introduction est disponible à l'adresse suivante : http://josephsalmon.eu/index.php?page=teaching_18_19&lang=fr, cours HLMA310 "logiciels scientifiques" Enfin une référence contenant la plupart des algorithmes usuels en informatique est le livre [CLRS01].

Description du cours

Après un bref aperçu de statistiques descriptives, ce cours introduira les méthodes élémentaires de statistiques et les concepts principaux que sont les tests, les intervalles de confiance

et l'estimation ponctuelle. Nous aborderons avec des exemples applicatifs les problématiques des modèles linéaires ainsi que de l'analyse de variance.

Notation

La note de CC compte pour 40% de la note finale et l'examen compte lui pour 60% de la note finale. Le CC sera oublié si l'examen est meilleur, ainsi la formule de calcul de la note est donc :

$$\text{Note final} = \max(0.4 \times \text{CC} + 0.6 \times \text{EXAM}; \text{EXAM})$$

- CC : moyenne du contrôle de mi-parcours en amphi et du TP noté. Le contrôle de mi-parcours est un quiz (avec point négatif pour chaque mauvaise réponse, il est donc déconseillé de répondre au hasard).
- EXAM : examen final

Bonus

2 pt supplémentaires sur **la note finale** peuvent être obtenus pour toute contribution à l'amélioration des cours (présentations, codes, etc.) sous les contraintes suivantes.

Contraintes :

- on obtient 1 point par amélioration
- seule la première amélioration reçue est "rémunérée", les autres ne gagnent plus rien
- il faut déposer un fichier **.txt** (taille <10 ko) en créant une fiche dans la partie du Moodle intitulée "Bonus - Proposition d'amélioration"
- détailler précisément (ligne de code, page des présentations, etc.) l'amélioration proposée, ce qu'elle corrige et/ou améliore
- pour les fautes d'orthographe : proposer au minimum 5 corrections par contribution
- chaque élève ne peut gagner que 2 points maximum de cette manière

Thèmes des séances :

- Semaine 1 : Consignes et introduction (cours 23/01)
- Semaine 2 : Statistiques descriptives (cours 30/01)
- Semaine 3 : Échantillonnage aléatoire simple (cours 6/02) - TD1
- Semaine 4 : Modèle statistique (cours 13/02) - TD2
- Semaine 5 : Tests (20/02) - TD3
- Semaine 6 : Estimation de paramètre (27/02) - TP1
- Semaine 7 : CC : Contrôle Continu (13/03) - TD4
- Semaine 8 : Intervalle de confiance / tests cas gaussiens (20/03) - TP2
- Semaine 9 : Régression linéaire (27/03) - TD5
- Semaine 10 : Régression linéaire II (03/04) - TP3
- Semaine 11 : Analyse de la variance (10/04) - TD6

Semaine 12 : Analyse de la variance II (17/04) - TP noté

Livres et ressources pour l'introduction aux statistiques

- **Modern Statistics for Modern Biology** ; S. Holmes, W. Huber, 2018, [HH18] <http://web.stanford.edu/class/bios221/book/>
- **Statistique générale pour utilisateurs. Méthodologie**, J. Pagès, 2010, [Pag10]
- **Statistics in Action with R** <http://sia.webpopix.org/index.html>

Page du cours :

http://josephsalmon.eu/index.php?page=teaching_18_19&lang=fr

Références

- [CGDM01] M. Cottrell, V. Genon-Catalot, C. Duhamel, and T. Meyre. *Exercices de probabilités, Licence - Master - Écoles d'ingénieur*. Cassini, 3^e édition, 2001.
- [CLRS01] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to algorithms*. MIT press, 2001.
- [GS01] G. R. Grimmett and D. R. Stirzaker. *Probability and random processes*. Oxford University Press, New York, third edition, 2001.
- [Gv13] G. H. Golub and C. F. van Loan. *Matrix computations*. Johns Hopkins University Press, Baltimore, MD, fourth edition, 2013.
- [HH18] S. Holmes and W. Huber. *Modern statistics for modern biology*. Cambridge University Press, 2018.
- [HJ94] R. A. Horn and C. R. Johnson. *Topics in matrix analysis*. Cambridge University Press, Cambridge, 1994. Corrected reprint of the 1991 original.
- [Ouv07] J-Y. Ouvrard. *Probabilités : Tome 2, Licence - CAPES*. Enseignement des mathématiques. Cassini, 2 édition, 2007.
- [Ouv08] J-Y. Ouvrard. *Probabilités : Tome 1, Licence - CAPES*. Enseignement des mathématiques. Cassini, 2 édition, 2008.
- [Pag10] J. Pagès. *Statistique générale pour utilisateurs. Méthodologie*. Presses universitaires de Rennes, 2010.
- [Sch05] J. R. Schott. *Matrix analysis for statistics*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], second edition, 2005.
- [TB97] L. N. Trefethen and D. III Bau. *Numerical linear algebra*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.