

## Statistique (spécifique Magistère)

Course title – Intitulé du cours	Statistique
Level / Semester – Niveau / semestre	L3 / S2 parcours DU Magistère Economiste Statisticien
Teacher – Enseignant responsable	Ecole d'Economie de Toulouse
Other teacher(s) – Autre(s) enseignant(s)	Eve Leconte
Other teacher(s) – Autre(s) enseignant(s)	Joseph Hachem
Lecture Hours – Volume Horaire CM	27h
TA Hours – Volume horaire TD	TD en autonomie
TP Hours – Volume horaire TP	18h
Course Language – Langue du cours	Français
TA and/or TP Language – Langue des TD et/ou TP	Français

### **Teaching staff contacts – Coordonnées de l'équipe pédagogique :**

Eve Leconte : eve.leconte@tse-fr.eu, bureau T 218

Joseph Hachem : joseph.hachem@tse-fr.eu, bureau T 216

Les étudiants peuvent poser des questions à la fin du cours, par mail ou prendre rendez-vous par mail pour rencontrer les enseignants.

### **Course Objectives – Objectifs du cours :**

Ce cours de statistique appliquée comporte deux parties :

#### **- Tests d'hypothèses (statistique inférentielle) avec Eve Leconte**

Cette partie du cours reprend et approfondit le chapitre sur les tests d'hypothèses du cours de Statistique de la L2 Économie et Mathématiques et initie les étudiants à la pratique des tests avec le logiciel R. A l'issue de cette partie, l'étudiant devra pouvoir formuler des hypothèses statistiques pour répondre à un problème donné et savoir choisir le test statistique adéquat pour trancher entre deux hypothèses en fonction de la nature des variables et du problème posé. Seront notamment étudiés les tests permettant d'étudier la liaison entre deux variables. L'étudiant saura mettre en oeuvre ces tests sur des jeux de données réelles à l'aide du logiciel R.

#### **- Simulations stochastiques (avec Joseph Hachem)**

Cette partie initie à la simulation de variables aléatoires et à la méthode de Monte-Carlo. A l'issue de cette partie, l'étudiant aura à sa disposition plusieurs méthodes pour simuler une variable aléatoire de loi donnée ; la simulation par l'inverse généralisée et la méthode de rejet seront abordées. L'étudiant saura également simuler une chaîne de Markov et commencer à en décrire le comportement asymptotique dans des cas simples. Enfin, l'étudiant devra être capable d'approcher des intégrales via la méthode de Monte-Carlo. Toutes ces méthodes devront pouvoir être mises en oeuvre via le langage Python.

### **Prerequisites – Pré requis :**

Statistique descriptive, variables aléatoires discrètes et continues usuelles (loi, espérance, variance), convergence en loi, estimateurs empiriques de l'espérance et de la variance.

### **Practical information about the sessions – Modalités pratiques de gestion du cours :**

Les ordinateurs et tablettes sont acceptés en cours et en TP.

Nous attendons des étudiants une participation active en cours et en TP.

De plus, la plupart des exercices de TD seront donnés sur Moodle et seront à faire en autonomie en dehors des cours.

Des corrigés seront ensuite mis à disposition sur Moodle.

L'acceptation des arrivées tardives est laissée à l'appréciation des enseignants.

### **Grading system – Modalités d'évaluation :**

Le contrôle continu compte pour 50 % et l'examen final pour 50 %.

Le contrôle continu se compose de deux projets :

- pour la partie Tests d'hypothèses : un projet avec le logiciel R (donné en mars et à rendre pendant les vacances de Pâques),
- pour la partie Simulations stochastiques : un projet avec le logiciel Python (donné en mars et à rendre pendant les vacances de Pâques).

L'examen terminal est un examen écrit de 2h sans documents sur l'ensemble du cours et des TD.

### **Bibliography/references – Bibliographie/références :**

#### **Pour les prérequis :**

« Statistique pour économistes et gestionnaires ». Brigitte Tribout, Pearson Education

#### **Pour l'apprentissage du logiciel R :**

R peut être téléchargé gratuitement sur le site du CRAN : <https://cran.r-project.org/>. Plusieurs manuels en anglais y sont accessibles.

Un manuel en ligne en français : [https://cran.r-project.org/doc/contrib/Paradis-rdebuts\\_fr.pdf](https://cran.r-project.org/doc/contrib/Paradis-rdebuts_fr.pdf)

#### **Pour l'apprentissage du logiciel Python :**

Pour utiliser Python, il est nécessaire d'installer un interpréteur Python, ainsi qu'un environnement de développement (ou IDE en anglais). Un exemple d'IDE est Pyzo, qu'on pourra télécharger ici, ainsi qu'un interpréteur : <https://pyzo.org/start.html>

Pour une introduction à Python : <https://docs.python.org/3/tutorial/introduction.html>

Nous utiliserons surtout les bibliothèques Numpy (<https://numpy.org/devdocs/user/quickstart.html>) et matplotlib.pyplot (<https://matplotlib.org/stable/tutorials/introductory/pyplot.html#sphx-glr-tutorials-introductory-pyplot-py>).

#### **Pour la partie du cours Tests d'hypothèses :**

« Probabilités, analyse des données et statistique ». Gilbert Saporta, Technip. Juillet 2011 (version révisée).

#### **Pour la partie du cours Simulations stochastiques :**

« python proba stat », Vincent Vigon, Independently published, octobre 2018.

### **Session planning – Planification des séances**

Chaque semaine, il y aura un cours sur la partie Tests d'hypothèses et un cours sur la partie Simulations stochastiques, ainsi qu'une séance de TP pour chaque partie, avec le logiciel R pour la partie Tests d'hypothèses et avec le logiciel Python pour la partie Simulations stochastiques.

#### **Plan du cours :**

##### **Partie Tests d'hypothèses**

- Rappels sur la théorie des tests d'hypothèses
- Tests d'ajustement du khi-deux de Pearson et de Kolmogorov-Smirnov
- Test du khi-deux d'indépendance
- Tests pour un échantillon (paramétriques et non paramétriques)
- Tests pour deux échantillons (comparaisons de variances, de moyennes et de médianes) paramétriques et non paramétriques
- Test de comparaison de plusieurs moyennes (test de l'analyse de variance) et de plusieurs médianes (test de Kruskal-Wallis)

##### **Partie Simulations stochastiques :**

- Rappels de probabilités : critères de convergence en loi, fonctions de répartition et notion d'inverse généralisée.
- Simulation par l'inverse généralisée, simulation des lois discrètes
- Simulation d'une loi normale : méthode de Box-Muller
- Simulation par la méthode de rejet, simulation de lois conditionnelles
- Introduction aux chaînes de Markov et simulation

-Méthode de Monte-Carlo pour l'approximation d'intégrales : précision et vitesse de convergence, comparaison avec les méthodes déterministes

**Distance learning – Enseignement à distance :**

Si la situation sanitaire l'exige, les enseignements pourront avoir lieu à distance sous forme de classe en ligne interactive, QCM et exercices en ligne, TP à distance, forum sous Moodle.