

Année universitaire 2024/2025

# Ingénierie Statistique et Financière - 2e année de Master en apprentissage

**Crédits ECTS : 60**

## LES OBJECTIFS DE LA FORMATION

Cette formation offre aux étudiants une formation solide leur permettant une insertion professionnelle rapide dans les métiers de l'industrie des services, de la finance, de l'assurance et de la statistique. Elle forme les étudiants aux instruments et aux raisonnements financiers, à la modélisation et aux statistiques, ainsi qu'au contexte réglementaire national, européen et international pour les banques, les assurances, et les sociétés de gestion.

- Maîtriser des méthodes quantitatives, la modélisation mathématiques et statistique, l'outil informatique et les principaux modèles d'évaluations des produits dérivés
- Analyser un problème, proposer et conduire à son terme une solution, en prenant en charge le traitement numérique et informatique ainsi que les aspects réglementaires, opérationnels
- Former aux techniques spécifiques de l'industrie des services (études économiques, marketing, gestion de la production, contrôle de la qualité, finance, assurance, etc.), et aux méthodes récentes en statistiques et Big Data
- Préparer l'étudiant aux problématiques des entreprises notamment dans le domaine de l'évaluation et de la mesure des risques et lui offrir l'accès à des postes à responsabilité
- Acquérir de solides compétences en matière de mathématiques financières et en analyse des données (Big Data).
- Acquérir une culture financière nécessaire à l'obtention de certaines certifications professionnelles (CFA, FCA etc...)

## PRÉ-REQUIS OBLIGATOIRES

- Titulaires d'un diplôme BAC+3 (180 crédits ECTS) ou équivalent à Dauphine, d'une université, d'une école d'ingénieurs ou d'un autre établissement de l'enseignement supérieur dans les domaines suivants : science, mathématiques appliquées, actuariat, probabilités et statistiques, gestion de portefeuille
- Etudiants issus d'une Grande Ecole avec de bonnes bases en statistique et finance (pour la voie finance)
- Avoir des connaissances suffisantes en statistique/finance
- Bonne maîtrise du calcul actuariel
- Les bases des principaux modèles d'évaluations des produits dérivés

## POURSUITE D'ÉTUDES

Plus de 80% des étudiants rejoignent des banques, sociétés d'assurance, conseil ou services financiers, tandis que les autres font le choix d'un VIE ou plus rarement d'une formation complémentaire.

Les métiers ciblés par cette formation sont :

- Chargé d'études statistiques, ingénieur statistique, chargé d'études marketing, consultant scoring, analyste financier, ingénieur financier, contrôleur des risques financiers, 'quant', trader, structureur, auditeur, ingénieur en modélisation financière, gérant de portefeuille, gestion actif/passif, ingénieur actuariaire
- contrôleur des risques de marché ou de crédit dans une banque d'investissement ou une société de gestion de portefeuille (Société Générale, BNP, Amundi...etc).
- Structureur ou analyste quantitatif dans une banque.
- Gérant ou assistant de gestion dans un hedge fund ou une société de gestion (Natexis Asset Management, Amundi...etc).
- Contrôleur de gestion à la direction financière d'une assurance (Axa, Groupama, Allianz...etc),
- gestionnaire des risques financiers dans une direction financière (banque, assurance, grande entreprise).
- Modélisateur, statisticien ou tarificateur pour un assureur ou une entreprise, consultant dans un cabinet de conseils (Ernst & Young, Deloitte, KPMG...etc) ou pour un éditeur de logiciels financiers (Murex, Bloomberg, Riskdata ...etc).

## PROGRAMME DE LA FORMATION

- Semestre 3
  - UE fondamentales S3
    - Processus Stochastiques
    - Méthodes actuarielles
    - Introduction à l'assurance vie et non vie
    - Solvabilité II
    - Méthodologie en gestion globale des Risques : VAR
    - Anglais des affaires
    - Introduction à l'apprentissage supervisé
    - Méthodes pour les modèles de régression
    - Introduction au Machine learning
    - Deep learning
    - Decentralized et Crypto Finance : new era of financial services
    - SAS, R et Python
  - UE complémentaires voie QRF S3
    - Statistiques et dynamique des produits dérivés
    - Modélisation stochastique du risque de crédit
- Semestre 4
  - UE fondamentales S4
    - Pratique des options
    - Culture Financière et pratique de Bloomberg
    - Python et pratique de la Data Science
  - UE complémentaires voie QRF S4
    - Modélisation stochastique des courbes de taux
    - Implémentation de modèles multivariés en finance et assurance
  - UE complémentaires voie MDB S4
    - Recent Advances in Data Sciences
    - Data Science for Business
    - Renforcement Learning
    - Machine Learning, Transformer et NLP
  - Conduite de projets et mémoire
    - Conduite de projet et Mémoire

## DESCRIPTION DE CHAQUE ENSEIGNEMENT

### Anglais des affaires

**ECTS** : 2

**Volume horaire** : 21

**Description du contenu de l'enseignement :**

1. Présentation des concepts et outils utilisés en management de projet, illustrée par des exemples concrets portant sur des projets, notamment dans le domaine de la Data Science.
2. Réalisation en groupe d'un projet de communication.

**Compétence à acquérir :**

Amener les étudiants à développer des stratégies qui leur permettent d'améliorer leurs compétences langagières, à l'écrit comme à l'oral. Un contenu lié à la recherche d'emploi et au monde du travail est abordé au moyen de simulations et

d'exercices de compréhension, de production et d'écoute.

---

## Conduite de projet et Mémoire

**ECTS** : 15

### Description du contenu de l'enseignement :

1. Présentation des concepts et outils utilisés en management de projet, illustrée par des exemples concrets portant sur des projets, notamment dans le domaine de la Data Science.
2. Réaliser en groupe un projet de communication.
3. Exposer en public son mémoire d'apprentissage

### Compétence à acquérir :

Familiariser les étudiants aux méthodes de communication dans le cadre d'un projet concret et leur apprendre les bases de la communication en entreprise (oral et écrit). Suivre le mémoire d'apprentissage

---

## Culture Financière et pratique de Bloomberg

**ECTS** : 2

**Volume horaire** : 15

### Description du contenu de l'enseignement :

1. Le bilan d'une entreprise et les différentes catégories de titres
2. Les actions, la dette, la dette hybride, le Tier 1, le Tier 2, le Tier 3
3. Les obligations convertibles, les mandatory convertibles
4. Les opérations en capital, les augmentations de capital, les FRESHs ; les Cocos
5. Les rachats d'actions, simples ou structurés
6. Les dividendes cash ou scrip, formules d'ajustements des dérivés
7. Les activités ECM, DCM, EQL, M&A
8. Séances pratiques sur Bloomberg :
  - construction de tableaux de bords en temps réel (BDP)
  - analyses historiques (BDH)
  - analyse financière
  - spreadsheets et templates Bloomberg
9. Mini projet de gestion de portefeuille

### Compétence à acquérir :

Connaitre les principales notions de corporate finance, analyser des données financières sur Bloomberg, faire un mini projet en gestion de portefeuille

---

## Data Science for Business

**ECTS** : 2

**Volume horaire** : 21

---

## Decentralized et Crypto Finance : new era of financial services

**ECTS** : 2

**Volume horaire** : 15

---

## Deep learning

**ECTS** : 2

**Volume horaire** : 18

### Description du contenu de l'enseignement :

- 1/ Deep learning : applications majeures, références, culture
- 2/ Types d'approches: supervisé, renforcement, non-supervisé

- 3/ Réseaux neuronaux: présentation des objets: neurones, opérations, fonction loss, optimisation, architecture
- 4/ Focus sur les algorithmes d'optimisation stochastique et preuve de convergence de SGD
- 5/ Réseaux convolutifs (CNN) : filtres, couches, architectures
- 6/ Technique: back-propagation, régularisation, hyperparamètres
- 7/ Réseaux pour des séquences : RNN, LSTM, Attention, Transformer
- 8/ réseaux génératifs (GAN, VAE)
- 9/ Environnements de programmation pour réseaux neuronaux: Tensorflow, Keras, PyTorch et travail sur les exemples vus en cours
- 10/ Si le temps permet: Stable diffusion, LLM

**Compétence à acquérir :**

introduction au deep learning

**Bibliographie, lectures recommandées :**

<https://turinici.com>

---

## Implémentation de modèles multivariés en finance et assurance

**ECTS :** 2

**Volume horaire :** 18

---

## Introduction au Machine learning

**ECTS :** 2

**Volume horaire :** 15

**Description du contenu de l'enseignement :**

1. Supervised and unsupervised learning
2. Calibration versus prediction: how to avoid over-fitting
3. Measure of the complexity of a model according to Vapnik-Chervonenkis
4. Vapnik-Chervonenkis's inequality and the control of the prediction error
5. Maximum margin SVMs and Gap tolerant classifiers
6. C-SVMs and duality
7. SVMs with kernels and Mercer's theorem
8. The simplex case
9. Mu-SVM, duality and reduced convex envelopes
10. Single class SVMs, anomaly detections and clustering
11. An introduction to Bootstrap, decision trees and random forests
12. Ridge Regression, penalization, and yield curve smoothing
13. The Representer theorem, Lasso, parsimony and duality.

**Compétence à acquérir :**

Théorie du statistical learning. Comprendre comment utiliser les Supports Vectors Machines pour l'apprentissage supervisé et non supervisé. Quelques application des méthodes de regressions pénalisées. Application à des problèmes de crédit et de courbe des taux.

**Mode de contrôle des connaissances :**

Examen

**Bibliographie, lectures recommandées :**

- [1] Pierre Brugiere: <https://hal.archives-ouvertes.fr/cel-01390383v2>
- [2] Wolfgang Karl Härdle, Rouslan Moro, Linda Hoffmann : Learning Machines Supporting Bankruptcy Prediction, SFB 649 Discussion Paper 2010-032
- [3] Dave DeBarr and Harry Wechsle: Fraud Detection Using Reputation Features SVMs, and Random Forests
- [4] Trevor Hastie, Robert Tibshirani, Jerome Friedman: The Elements of Statistical Learning
- [5] Christopher Bishop: Pattern Recognition and Machine Learning

## Introduction à l'apprentissage supervisé

**ECTS** : 3

**Volume horaire** : 21

### Description du contenu de l'enseignement :

- 1- Analyse factorielle discriminante
- 2- Analyse discriminante linéaire et quadratique
- 3- Classification bayésienne à l'aide de modèles de mélange
- 4- Classifieur bayésien et classifieur bayésien naïf
- 5- Sélection de modèles de mélange parcimonieux
- 6- Arbres de décision
- 7- Forêts aléatoires

L'ensemble de ces méthodes enseignées est illustré par des démonstrations du logiciel R sur des jeux de données réel (principalement Analyse Discriminante linéaire et quadratique, Classification bayésienne gaussienne, Classifieur bayésien naïf, Forêts aléatoires).

### Compétence à acquérir :

Ce cours présente les méthodes élémentaires d'apprentissage supervisé suivantes : analyse factorielle discriminante, classification bayésienne à l'aide de modèles de mélange, arbres de décision et forêts aléatoires. Les propriétés théoriques et différentes formulations de ces méthodes sont présentées. Leurs mises en oeuvre, ainsi que celles de leurs variantes, sont illustrées à l'aide de traitements de données effectués avec le logiciel R. L'objectif de ce cours est l'acquisition de la maîtrise de ces méthodes élémentaires d'apprentissage supervisé.

### Mode de contrôle des connaissances :

Examen

### Bibliographie, lectures recommandées :

- Benzecri, J.-P. (1980) Pratique de l'analyse des données. Dunod. Paris.
- Bouveyron, C., Celeux, G., Murphy, T., & Raftery, A. (2019) Model-Based Clustering and Classification for Data Science: With Applications in R, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge: Cambridge University Press.
- Breiman, L., Friedman, J.H., Olshen, R., and Stone, C.J. (1984). Classification and Regression Trees, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific California.
- Hastie, T., Tibshirani, R., Friedman, J. (2009) The Elements of Statistical Learning : Data Mining, Inference, and Prediction, Second Edition, Springer Series in Statistics.
- James, G., Witten, D., Hastie, T., Tibshirani, R., Taylor, J., (2023) An Introduction to Statistical Learning: With Applications in Python, Springer International Publishing.
- Lebart, L., Piron, M., Morineau, A. (2006) (4ème édition, refondue) Statistique Exploratoire Multidimensionnelle, 480 pages, Dunod.
- Saporta, G. (2006), Probabilités, Analyse des données et Statistique, 656 pages, Technip.

---

## Introduction à l'assurance vie et non vie

**ECTS** : 2

**Volume horaire** : 21

### Description du contenu de l'enseignement :

1. Définir les termes et les acteurs d'une opération d'assurance
2. Donner des éléments statistiques sur le secteur de l'assurance
3. Rappeler des éléments de probabilité et de mathématiques financières
4. Déterminer la tarification des engagements vie et non vie
5. Exposer les formes de garanties proposées en vie et en non vie
6. Calculer les engagements techniques des contrats d'assurance
7. Étudier la gestion du risque au niveau de l'organisme assureur

8. Donner des notions de comptabilité et de réglementation propre à l'assurance
9. Présenter les principes de la réassurance
10. Étudier la notion de solvabilité d'un organisme assureur et quelques éléments prudentiels

**Compétence à acquérir :**

Présenter les principaux modèles de l'assurance vie et non vie.

---

## Machine Learning, Transformer et NLP

ECTS : 2

Volume horaire : 18

---

## Modélisation stochastique des courbes de taux

ECTS : 3

Volume horaire : 21

**Description du contenu de l'enseignement :**

1. Quelques outils de calcul stochastique : rappels
2. Généralités sur les taux d'intérêt
3. Produits de taux classiques
4. Modèle LGM à un facteur
5. Modèle BGM (Brace, Gatarek et Musiela) / Jamishidian
6. Modèles à volatilité stochastique

**Compétence à acquérir :**

Ce cours est consacré aux modèles de taux d'intérêts à temps continu. Au travers de nombreux exemples, on décrira leurs utilisations pour évaluer les produits dérivés sur taux d'intérêt.

---

## Modélisation stochastique du risque de crédit

ECTS : 2

Volume horaire : 21

---

## Méthodes actuarielles

ECTS : 3

Volume horaire : 21

**Description du contenu de l'enseignement :**

1. Instruments et marchés (marchés monétaires, marchés obligataires)
2. Mesure et couverture du risque de taux (duration, convexité, ACP)
3. Reconstitution de la structure par terme des taux (modèles à splines, modèles paramétriques)
4. Théories de la structure par terme des taux (anticipations pures, prime de risque pure, segmentation, anticipations biaisées)
5. Gestion passive (tracking error, échantillonnage stratifié)
6. Gestion active (roll-down, barbell, bullet, butterfly)
7. Produits dérivés de taux (futures & swaps)

**Compétence à acquérir :**

Fournir une explication détaillée de la structure par terme des taux et apporter un éclairage sur les différentes stratégies de gestion et leur mise en œuvre.

---

## Méthodes pour les modèles de régression

ECTS : 3

Volume horaire : 21

**Description du contenu de l'enseignement :**

L'objectif de ce cours est de présenter aux étudiants des connaissances fondamentales, sur la régression d'un point de vue

théorique ainsi que sur le code lié à ce domaine.

- Rappels théoriques sur le modèle linéaire gaussien multivarié, Anova, Ancova, sélection de modèle, validation du modèle...
- Étude et traitement des outliers en régression.
- Étude des différents critères (AIC, BIC, Cp-Mallows,...) et sélection de modèles.
- Analyse complète de différents modèles linéaires gaussiens multivariés sous R à partir de jeux de données réelles.
- Estimateurs biaisés (Lasso, Ridge, Elastic-Net, PLS,...)
- Performance de généralisation (PRESS sur échantillon tests,...), Validation Croisée,...
- Comparaison des différentes procédures (ML, Lasso,...) sous R à partir de jeux de données réelles.
- Modèles linéaires généralisés (régression poissonnière, régression logistique,...).
- Régression logistique d'un point de vue théorique et sous forme de TP avec des données réelles : déclaration du modèle, validation du modèle, sélection de modèle, odd ratio, matrice de confusion courbe ROC, AUC.

**Compétence à acquérir :**

A la suite de ce module, les étudiants seront capables de comprendre la régression d'un point de vue théorique et de coder les différentes procédures étudiées. Ils auront le recul nécessaire pour préselectionner des procédures adaptées à la spécificité du jeu de données et sélectionner celles ayant les meilleures performances de généralisation.

**Mode de contrôle des connaissances :**

Examen

---

## Méthodologie en gestion globale des Risques : VAR

ECTS : 3

Volume horaire : 21

**Description du contenu de l'enseignement :**

1. Introduction et définition de la Value at Risk
2. Méthodes et méthodologies de calcul
3. Choix de distribution de probabilité pour positions optionnelles
4. Mesure de risque de marché et RiskMetrics
5. Risque de crédit et exigences réglementaires
6. Risque de corrélation défavorable, liquidité et xVA
7. Expected Shortfall et VaR sur Valeurs extrêmes

**Compétence à acquérir :**

Clarifier la notion de risque et présenter les principales techniques et méthodes de VaR permettant de mesurer, analyser et prédire le risque.

Le risque de marché fera l'objet d'une attention particulière au travers de l'analyse de la VaR. Les méthodes de gestion globale du risque de marché lorsque les sources d'incertitudes sont multiples seront également étudiées.

---

## Pratique des options

ECTS : 2

Volume horaire : 15

**Description du contenu de l'enseignement :**

1. Mise en situation concrète du métier de trading (market making)
2. Pricing des options complexes à partir de celui des options vanilles
3. Les risques dans la vraie vie
4. Au-delà des grecques

**Compétence à acquérir :**

- Comprendre les responsabilités d'un market maker d'options
- Maîtriser les implications concrètes au-delà des équations de la gestion d'un portefeuille d'options
- Acquérir des réflexes afin de repérer rapidement les principales sources de risques

---

## Processus Stochastiques

ECTS : 3

**Volume horaire** : 30

**Description du contenu de l'enseignement :**

1. Intégrale stochastique
2. EDS et théorèmes de représentation
3. EDSR et théorèmes de représentation
4. Application au contrôle stochastique

**Compétence à acquérir :**

Approfondir les notions de processus stochastiques, équations différentielles stochastiques progressives (EDS) et rétrogrades (EDSR), lien avec les équations aux dérivées partielles (e.d.p) et application au contrôle stochastique

**Mode de contrôle des connaissances :**

Examen

---

## Python et pratique de la Data Science

**ECTS** : 2

**Volume horaire** : 15

**Description du contenu de l'enseignement :**

- Préparation des données, EDA. Traitement des classes déséquilibrées
  - Les Réseaux de neurones ANN et leur optimisation avec les Hyperparametres
  - Les Réseaux de neurones RNN et CNN
  - L'apprentissage non supervisé avec KNN et les Autoencodeurs
  - Le modèle XGBoost
  - Le cours se déroule d'une manière générale en 2 parties: Rappels théoriques et approfondissements, puis projet à mener en trinôme.
- 

## Recent Advances in Data Sciences

**ECTS** : 3

**Volume horaire** : 21

**Description du contenu de l'enseignement :**

Le domaine du Machine Learning et du Deep Learning évoluant sans cesse plus rapidement, il est essentiel d'avoir des bases solides dans ces deux domaines pour naviguer dans les nombreux articles de recherche du domaine. Nous traiterons de notions réinterprétés ou présentés entre 2017 et 2023 à la lumière d'articles plus anciens.

Les séances seront organisées par thème :

1. Bon départ d'un réseau de neurones
2. Calibration en Machine Learning
3. Mise à jour moderne des poids d'un réseau de neurones
4. Avancées des *Large Language Models* pour les réseaux de neurones en général
5. Tokenization et impacts

Pour chaque séances plusieurs TP avec PyTorch seront proposés pour manipuler le cours. Les séances restantes seront dédiées à la préparation d'une soutenance finale portant sur une ou plusieurs des notions abordées dans les 5 séances.

**Compétence à acquérir :**

- Lecture, implémentation et critique d'un papier de recherche en Machine Learning
- Utilisation de PyTorch et Latex
- Capacité de synthèse et de vulgarisation de notions complexes en Machine Learning
- Culture générale sur les domaines et les acteurs

**Mode de contrôle des connaissances :**

Soutenance d'un sujet de recherche proposé, avec rédaction d'un rapide compte-rendu, des slides et un notebook.

**Bibliographie, lectures recommandées :**

- **Machine Learning et Deep Learning**
  - [Hands-on Machine Learning with Scikit-Learn](#), Aurélien Géron, O'Reilly



- [Deep Learning avec TensorFlow](#), Aurélien Géron, O'Reilly
  - [Deep Learning with Python](#), François Chollet, Manning
  - **Culture générale**
    - [Weapons of Math Destruction](#), Cathy O'Neil, Crown Books
    - [Quand la machine apprend](#), Yann Le Cun, Odile Jacob
    - [De l'autre côté de la machine](#): Voyage d'une scientifique au pays des algorithmes, Aurélie Jean, Édition de l'observatoire
- 

## Renforcement Learning

**ECTS** : 2

**Volume horaire** : 21

### Description du contenu de l'enseignement :

- 1/ Introduction au renforcement learning
- 2/ Formalisme théorique : « Markov decision processes » (MDP), fonction valeur (équation de Bellman et Hamilton- Jacobi – Bellman) etc.
- 3/ Stratégies usuelles, sur l'exemple de "multi-armed bandit"
- 4/ Stratégies en deep learning: Q-learning et DQN
- 5/ Stratégies en deep learning: SARSA et variantes
- 6/ Stratégies en deep learning: Actor-Critic et variantes
- 7/ Implémentations Python variées
- 8/ Perspectives.

### Compétence à acquérir :

introduction au deep reinforcement learning, avec une vision machine learning empirique: principaux algorithmes, implementations pratiques (gym)

### Bibliographie, lectures recommandées :

<https://turinici.com>

---

## SAS, R et Python

**ECTS** : 3

**Volume horaire** : 21

---

## Solvabilité II

**ECTS** : 2

**Volume horaire** : 18

### Description du contenu de l'enseignement :

1. Introduction au contexte de solvabilité
2. Présentation du calcul de solvabilité
3. Dispositifs de gestion des risques (ORSA)
4. Analyse prospective & introduction à l'appétence aux risques

### Compétence à acquérir :

Fournir aux étudiants des connaissances sur le contrôle prudentiel des organismes d'assurances. Leur permettra d'appréhender la complexité des problèmes comptables et les mécanismes d'évaluation du ratio de solvabilité. Le fonctionnement et l'approche de la gestion des risques dans le secteur de l'assurance seront présentés dans ce nouveau contexte.

---

## Statistiques et dynamique des produits dérivés

**ECTS** : 2

**Volume horaire** : 21

### Description du contenu de l'enseignement :

- 1/ Approches en probabilité historique (gestion de portefeuille classique), portefeuilles optimaux, beta, arbitrage, APT

2/ Valuation de produits dérivés et probabilité risque neutre

3/ Trading de volatilité, volatilité locale et calibration

4/ Assurance du portefeuille: stop-loss, options, CPPI, CRP - Constant Mix

5/ Options exotiques ou cachées: ETF short, etc.

**Compétence à acquérir :**

approche pratique et empirique des produits dérivés et de la gestion de risques tout en se basant sur une formalisation stochastique avancée

**Bibliographie, lectures recommandées :**

Voir <https://turinici.com>

---

**Université Paris Dauphine - PSL - Place du Maréchal de Lattre de Tassigny - 75775 PARIS Cedex 16 - 21/11/2024**