

# STATISTIQUES (STA)

## Les cours en statistiques (STA) sont offerts par la Faculté des sciences

### STA 1391 Descriptive Statistics (3 units)

Topics from descriptive statistics: histograms and boxplot; average and observed standard deviation; elementary probability; normal distribution; statistical estimation and hypothesis testing; correlation and regression. Examples analyzed with statistical software.

**Course Component:** Discussion Group, Lecture

The courses ECO 2150, MAT 1371, MAT 2371, MAT 2377, MAT 2379, STA 1391, STA 2391, STA 2392, cannot be combined for units. This course cannot count for unit in any program in the Faculty of Science or Engineering. Previously MAT 1371.

### STA 1791 Statistique descriptive (3 crédits)

Des éléments de statistiques descriptives : histogrammes et diagramme à boîte et moustaches; moyenne et écart-type de l'échantillon; probabilités élémentaires; la loi normale; l'estimation statistique et tests d'hypothèses; corrélation et régression. Exemples analysés avec un logiciel statistique.

**Volet :** Groupe de discussion, Cours magistral

Les cours ECO 2550, MAT 1771, 2771, 2777, 2779, STA 1791, 2791 et 2792 ne peuvent être combinés pour l'obtention de crédits. Ce cours ne compte pas pour des crédits dans un programme de la Faculté des sciences ou de génie.. Antérieurement MAT 1771.

### STA 2100 Introduction to Statistics (3 units)

Introduction to the theory of statistical inference. Parametric point and interval estimation. Maximum likelihood estimation. Properties of estimators. Principles of hypothesis testing. Confidence intervals and tests for means and proportions (one and two sample). Introduction to linear models. Use of statistical software.

**Course Component:** Laboratory, Lecture

Prerequisite: MAT 2371. The courses ADM 2303, ECO 2150, HSS 2381, MAT 2375, MAT 2377, MAT 2379, PSY 2106, STA 2100, STA 2391, STA 2392 cannot be combined for units. Previously MAT 2375.

### STA 2391 Probability and Statistics for Engineers (3 units)

A concise survey of: combinatorial analysis; probability and random variables; discrete and continuous densities and distribution functions; expectation and variance; normal (Gaussian), binomial and Poisson distributions; statistical estimation and hypothesis testing; method of least squares, correlation and regression. The emphasis is on statistics and quality control methods for engineers. Previously MAT 2377.

**Course Component:** Lecture

Prerequisite: MAT 1320 or MAT 1321 or MAT 1327 or MAT 1330. The course MAT 1322 or MAT 1332 is corequisite to STA 2391. ADM 2303, ECO 2150, HSS 2381, MAT 1371–1372, 2371, 2375, 2377, 2379, PSY 2106, STA 1391, 2100, 2391–2392 cannot be combined for units.

### STA 2392 Introduction to Biostatistics (3 units)

Descriptive statistics using a software package. A concise survey of probability. The normal distribution. The central limit theorem and statistical estimation illustrated via simulation. Hypothesis testing, the design of experiments, paired sampling, categorical data and regression. Examples from the biosciences analyzed with statistical software.

**Course Component:** Lecture

Prerequisite: MAT 1320 or MAT 1321 or MAT 1327 or MAT 1330.

ADM 2303, ECO 2150, HSS 2381, MAT 1371, MAT 1372, MAT 1373, MAT 2375, MAT 2377, MAT 2379, PSY 2106, STA 1391, STA 2100, STA 2391, STA 2392 cannot be combined for units. Previously MAT 2379.

### STA 2500 Introduction à la statistique (3 crédits)

Introduction à la théorie de l'inférence statistique. Estimation paramétrique ponctuelle et par intervalle. Estimation par maximum de vraisemblance. Propriétés des estimateurs. Principes de tests d'hypothèses. Intervalles de confiance et tests d'hypothèses pour des moyennes et proportions (un et deux échantillon). Introduction aux modèles linéaires. Utilisation de logiciels de statistique.

**Volet :** Laboratoire, Cours magistral

Préalable : MAT 2771. Les cours ADM 2703, ECO 2550, HSS 2781, MAT 2775, MAT 2777, MAT 2779, PSY 2506, STA 2500, STA 2791, STA 2792 ne peuvent être combinés pour l'obtention de crédits. Antérieurement MAT 2775.

### STA 2791 Probabilités et statistique pour ingénieurs (3 crédits)

Brève introduction aux sujets suivants: analyse combinatoire, probabilités et variables aléatoires, fonctions de densité et de répartition pour les variables de type discret ou continu. Espérance mathématique et variance; lois normale, binomiale et de Poisson; estimation et tests d'hypothèses; méthode des moindres carrés, corrélation et régression. La théorie est illustrée par des simulations. Le contenu du cours est orienté surtout vers la statistique et les méthodes de contrôle de la qualité pour les ingénieurs. Les cours ADM 2703, ECO 2550, HSS 2781, MAT 1771, 1772, 2771, 2775, 2777, 2779, PSY 2506, STA 1791, 2500, 2791 et 2792 ne peuvent être combinés pour les crédits. Antérieurement MAT 2777.

**Volet :** Cours magistral

Préalable : MAT 1720 ou MAT 1721 ou MAT 1727 ou MAT 1730. Le cours MAT 1722 ou MAT 1725 ou MAT 1732 est concomitant à STA 2791.

### STA 2792 Introduction à la biostatistique (3 crédits)

La statistique descriptive via un logiciel statistique. Introduction à la théorie des probabilités. La densité normale. Le théorème limite central et l'estimation statistique via simulation. Les tests d'hypothèse, plans d'expérience, comparaison de deux échantillons, analyse des données catégoriques et régression. Les exemples en sciences biologiques sont analysés par un logiciel statistique. Les cours ADM 2703, ECO 2550, HSS 2781, MAT 1771, MAT 1772, MAT 1773, MAT 2775, MAT 2777, MAT 2779, PSY 2506, STA 1791, STA 2500, STA 2791, STA 2792 ne peuvent être combinés pour l'obtention de crédits. Antérieurement MAT 2779.

**Volet :** Cours magistral

Préalable : MAT 1720 ou MAT 1721 ou MAT 1727 ou MAT 1730. ADM 2703, ECO 2550, HSS 2781, MAT 1771, 1772, 1773, 2775, 2777, 2779, PSY 2506, STA 1791, 2500, 2791 et 2792 ne peuvent être combinés pour les crédits. Antérieurement MAT 2779.

### **STA 3100 Introduction to Mathematical Statistics (3 units)**

Theory of statistical methods at an advanced mathematical level. Probability and limit theorems pertaining to statistical inference. Distribution theory and sampling distributions. Principles of parametric estimation. Likelihood-based inference and hypothesis tests. Introduction to non-parametric inference.

**Course Component:** Lecture

Prerequisites: MAT 3172, (MAT 3375 or STA 3300). The courses MAT 3175 and STA 3100 cannot be combined for units. Previously MAT 3175.

### **STA 3300 Regression Analysis (3 units)**

Modeling relationships between outcome (dependent) variables and covariates (independent variables) using simple and multiple linear regression models. Estimation and hypothesis testing using least squares and likelihood methods. Performing model diagnostics and assessing goodness of fit properties. Variable selection and finding the best fit. Non-linear regression and transformations. Weighted regression and generalized least square. Analysis of data using statistical software packages.

**Course Component:** Laboratory, Lecture

Prerequisites: MAT 1341, MAT 2371, (MAT 2375 or MAT 2379 or STA 2100 or STA 2392). The courses MAT 3375 and STA 3300 cannot be combined for units. Previously MAT 3375.

### **STA 3301 Analysis of Experimental Designs (3 units)**

Theory and analysis of common experimental designs: factorial designs, designs with blocking, split-plot, random effects models. Sample size and power determination. Simultaneous inference. Nested and crossed factors. Introduction to other designs: fractional factorial, response surface designs. Analysis of data using statistical software packages.

**Course Component:** Laboratory, Lecture

Prerequisites: MAT 1341, MAT 2371, (MAT 3375 or STA 3300). The courses MAT 3378 and STA 3301 cannot be combined for units. Previously MAT 3378.

### **STA 3302 Introduction to Time Series Analysis (3 units)**

Time domain methods: detrending, dealing with non stationarity and seasonality; autoregressive moving average (ARMA) models: forecasting, estimation, diagnostics. Selected topics from: state space methodology, financial time series, time series regression, spectral domain. Time series data will be analyzed using software packages.

**Course Component:** Laboratory, Lecture

Prerequisites: MAT 2371, (MAT 2375 or MAT 2379 or STA 2100 or STA 2392). The courses MAT 3379 and STA 3302 cannot be combined for units. Previously MAT 3379.

### **STA 3500 Introduction à la statistique mathématique (3 crédits)**

Théorie des méthodes statistiques couverte à un niveau de mathématique avancé. Probabilité et les théorèmes limites liés à l'inférence statistique. Distributions d'échantillonnage. Principes de l'estimation paramétrique. Inférence et tests d'hypothèses basés sur la fonction de vraisemblance. Introduction aux méthodes non-paramétriques.

**Volet :** Cours magistral

Préalables : MAT 3572, (MAT 3775 ou STA 3700). Les cours MAT 3575 et STA 3500 ne peuvent être combinés pour l'obtention de crédits. Antérieurement MAT 3575.

### **STA 3700 Analyse de la régression (3 crédits)**

Construction de modèles pour la relation entre une variable réponse (variable dépendante) et des variables explicatives (variables indépendantes) au moyen de la régression linéaire simple et multiple. Estimation et tests d'hypothèses par la méthode des moindres carrés et méthode de vraisemblance. Diagnostics de modèle et évaluation de la qualité de l'ajustement. Régression non-linéaire et transformations. Régression pondérée et méthode des moindres carrés généralisées. Analyse de données à l'aide de logiciels de statistique.

**Volet :** Laboratoire, Cours magistral

Préalables : MAT 1741, MAT 2771, (MAT 2775 ou MAT 2779 ou STA 2500 ou STA 2792). Les cours MAT 3775 et STA 3700 ne peuvent être combinés pour l'obtention de crédits. Antérieurement MAT 3775.

### **STA 3701 Analyse des plans d'expérience (3 crédits)**

Théorie et analyse des plans d'expérience suivants: plans factoriels, plans en blocs randomisés, plans à parcelles divisées, modèle à effets aléatoires. Taille de l'échantillon et calcul de la fonction de puissance. Inférence simultanée. Facteurs croisés et emboîtés. Introduction aux plans fractionnels, surface d'une fonction de régression. Analyse de données à l'aide de logiciels de statistique.

**Volet :** Laboratoire, Cours magistral

Préalables : MAT 1741, MAT 2771, (MAT 3775 ou STA 3700). Les cours MAT 3778 et STA 3301 ne peuvent être combinés pour l'obtention de crédits. Antérieurement MAT 3778.

### **STA 3702 Introduction aux séries chronologiques (3 crédits)**

Approche temporelle : décomposition des modèles de tendance, processus non-stationnaire et saisonnalité, modèles autorégressifs et de moyennes mobiles (ARMA): prédiction, estimation, évaluation des modèles. Sujets choisis parmi les suivants: modèles à espace d'état, séries chronologiques financières, régression temporelle, approche spectrale. Analyse de données de séries chronologiques à l'aide de logiciels de statistique.

**Volet :** Laboratoire, Cours magistral

Préalables : MAT 2771, (MAT 2775 ou MAT 2779, ou STA 2500 ou STA 2792). Les cours MAT 3779 et STA 3302 ne peuvent être combinés pour l'obtention de crédits. Antérieurement MAT 3779.

### **STA 4301 Bayesian Inference (3 units)**

An introduction to the theory and practice of modern Bayesian inference. Choice of prior distributions and calculation of the posterior distribution for single and multi-parameter models. Computational approaches to Bayesian inference such as Markov Chain Monte Carlo and Gibbs sampling. Hierarchical and regression modelling in a Bayesian framework. Use of statistical software.

**Course Component:** Lecture

Prerequisites: MAT 2342, (MAT 3375 or STA 3300). The courses MAT 4381 and STA 4301 cannot be combined for units. Previously MAT 4381.

### **STA 4302 Advanced Regression (3 units)**

Topics in linear regression: influential values, high leverage points, model selection, model comparison. Regularization methods: ridge regression, lasso regression. Robust regression: outlier, breakdown point, influence function, M-estimators. Non-parametric regression and smoothing methods: kernel regression, local polynomials, smoothing splines. Additional topics at instructor's discretion.

**Course Component:** Lecture

Prerequisite: MAT 3375 or STA 3300. The courses MAT 4380 and STA 4302 cannot be combined for units. Previously MAT 4380.

### **STA 4303 Categorical Data Analysis (3 units)**

Analysis of multi-way contingency tables. Tests for proportions: dependent, independent, trend. Measures of risk and association. Study designs leading to categorical data. Logistic models with binary response. Log-linear models for multi-way tables. Additional topics: polytomous and ordinal logistic regression. Analysis with statistical software packages.

**Course Component:** Lecture

Prerequisite: MAT 3375 or STA 3300. The courses MAT 4378 and STA 4303 cannot be combined for units. Previously MAT 4378.

### **STA 4304 Generalized Linear Models (3 units)**

Exponential family of distributions. Likelihood ratio, Wald and score tests. Model building and assessment including: deviance, goodness-of-fit, prediction and classification, multi-collinearity. Logistic regression. Poisson regression, overdispersion. Log-linear models. Additional topics in generalized linear models chosen by the instructor. Analysis of data using statistical software packages.

**Course Component:** Lecture

Prerequisite: MAT 3375 or STA 3300. The courses MAT 4382 and STA 4304 cannot be combined for units. Previously MAT 4382.

### **STA 4305 Survey Sampling (3 units)**

Finite populations and finite population parameters, sampling design, simple random sampling, systematic sampling, Bernoulli sampling, stratified sampling, sample allocation, estimation of a domain total/mean, estimation of ratio, ratio estimator, post-stratified estimator, difference estimator, simple linear regression estimator, single-stage and two-stage cluster sampling, treatment of missing data. Analysis of data using statistical software packages.

**Course Component:** Lecture

Prerequisites: MAT 2371, (MAT 3375 or STA 3300). The courses MAT 3377, MAT 4379 and STA 4305 cannot be combined for units. Previously MAT 4379.

### **STA 4306 Computational Statistics (3 units)**

Computational workflow and notebooks. Integration with quadrature and Monte Carlo (rejection, importance and Markov chain sampling). Optimization techniques such as the EM algorithm and gradient descent. Error estimation such as bootstrap and cross-validation. Simulation studies. Statistical programming and the use of statistical software. Other topics in computational statistics at instructor's discretion.

**Course Component:** Lecture

Prerequisites: MAT 3172, (MAT 3373 or MAT 3375 or STA 3300). The courses MAT 4374 and STA 4306 cannot be combined for units. Previously MAT 4374.

### **STA 4307 Multivariate Statistical Methods (3 units)**

Random vectors and matrices. Multivariate normal distribution and its properties. Inference on multivariate normal distribution: estimation, hypothesis testing. Principal component analysis. Cluster analysis. Linear discriminant analysis. Canonical correlation analysis. Analysis of data using statistical software packages.

**Course Component:** Lecture

Prerequisites: (MAT 2141 or MAT 2342), (MAT 3375 or STA 3300). The courses MAT 4375 and STA 4307 cannot be combined for units. Previously MAT 4375.

### **STA 4320 Topics in Statistics (3 units)**

Selected topics in statistics.

**Course Component:** Lecture

Prerequisite: MAT 2375 or STA 2100 (additional prerequisites may be added depending on the topic). Previously MAT 4376.

### **STA 4340 Statistics Laboratory (3 units)**

This course provides students with the opportunity to apply the statistical techniques and theory learned in their program to design, collection, analysis, interpretation and reporting of real data. Through case studies, students will propose and implement solutions to realistic scientific problems. Course work will include class discussions, teamwork, research methods and reading scientific articles, performing exploratory data analysis, formulating statistical analysis plans, analyzing data, project management and reproducible workflows, interpretation of results from statistical analysis, writing statistical analysis reports, and oral presentations. Learning statistical consultation skills. Reserved for students enrolled in 4th year of the Honours B.Sc. in Statistics. The courses MAT 4383 and STA 4340 cannot be combined for units.

**Course Component:** Lecture

Prerequisites: (MAT 3175 or STA 3100), (MAT 3375 or STA 3300), and six course units from MAT 3378, MAT 3379, MAT 4374, MAT 4375, MAT 4378, MAT 4380, MAT 4382, STA 3301, STA 3302, STA 4302, STA 4303, STA 4304, STA 4306, STA 4307. Previously MAT 4383.

### **STA 4701 Inférence Bayésienne (3 crédits)**

Introduction à la théorie et la pratique de l'inférence bayésienne moderne. Choix de lois a priori et calcul de la loi a posteriori pour les modèles avec un et plusieurs paramètres. Approches computationnelles pour l'inférence bayésienne telle que les chaînes de Markov Monte Carlo et échantillonneur de Gibbs. Modèles hiérarchiques et régression linéaire dans un cadre bayésienne. Utilisation d'un logiciel statistique.

**Volet :** Cours magistral

Préalables : MAT 2742, (MAT 3775 ou STA 3700). MAT 4781 et STA 4701 ne peuvent être combinés pour l'obtention de crédits. Antérieurement MAT 4781.

### **STA 4702 Régression avancée (3 crédits)**

Sujets choisis en régression linéaire : observations influentes, points de levier, choix de modèles, comparaison de modèles. Méthodes de régularisation : régression ridge, régression lasso. Régression robuste : valeur aberrante, point de rupture, fonction d'influence, M-estimateurs. Régression non paramétrique et méthodes de lissage : régression à noyau, régression polynomiale, fonctions splines à lissage. Autres sujets choisis par le professeur.

**Volet :** Cours magistral

Préalable : MAT 3775 ou STA 3700. Les cours MAT 4780 et STA 4702 ne peuvent être combinés pour l'obtention de crédits. Antérieurement MAT 4780.

### **STA 4703 Analyse des données catégoriques (3 crédits)**

Analyse des tableaux de contingence à plusieurs facteurs. Tests d'hypothèses sur les proportions: échantillons dépendants, indépendants et test de tendance. Mesures de risque et d'association. Plans d'expérience conduisant aux données catégorielles. Modèle logistique pour les données binaires. Modèles log-linéaires pour des tableaux de contingences multidimensionnels. Autres sujets choisis parmi les suivants: régression logistique multinomiale, régression ordinale. Analyse des données à l'aide de logiciels de statistique.

**Volet :** Cours magistral

Préalable : MAT 3775 ou STA 3700. Les cours MAT 4778 et STA 4703 ne peuvent être combinés pour l'obtention de crédits. Antérieurement MAT 4778.

### **STA 4704 Modèles linéaires généralisés (3 crédits)**

Famille exponentielle. Test du rapport de vraisemblance, test de Wald et test du multiplicateur de Lagrange. Construction et évaluation des modèles y compris : déviance, l'ajustement, la prévision et la classification, multicollinéarité. Régression logistique. Régression de Poisson, sur-dispersion. Modèles log-linéaires. Autres sujets choisis par le professeur. Analyse de données à l'aide de logiciels de statistique.

**Volet :** Cours magistral

Préalable : MAT 3775 ou STA 3700. Les cours MAT 4782 et STA 4704 ne peuvent être combinés pour l'obtention de crédits. Antérieurement MAT 4782.

### **STA 4705 Échantillonnage (3 crédits)**

Populations finies et paramètres de populations finies, plan d'échantillonnage, échantillonnage aléatoire simple, échantillonnage systématique, échantillonnage de Bernoulli, échantillonnage stratifié, répartition de l'échantillon, estimation du total et d'une moyenne d'un domaine, estimation d'un ratio, estimateur par le ratio, estimateur post-stratifié, estimateur par la différence, estimateur par la régression linéaire simple, plans à un et à deux degrés, traitement des données manquantes. Analyse de données à l'aide de logiciels de statistique

**Volet :** Cours magistral

Préalables : MAT 2771, (MAT 3775 ou STA 3700). Les cours MAT 3777, MAT 4779 et STA 4705 ne peuvent être combinés pour l'obtention de crédits. Antérieurement MAT 4779.

### **STA 4706 Statistique computationnelle (3 crédits)**

Processus pour les tâches computationnelles. Intégration numérique et par Monte Carlo (méthode de rejet, échantillonnage préférentiel et applications aux chaînes de Markov Monte Carlo). Techniques d'optimisation telles que l'algorithme espérance-maximisation et l'algorithme du gradient. Estimation de l'erreur au moyen du bootstrap et de la validation croisée. Etudes par simulation. Programmation statistique et utilisation de logiciels de statistique. Autres sujets à la discrétion du professeur.

**Volet :** Cours magistral

Préalables : MAT 3572, (MAT 3773 ou MAT 3775 ou STA 3700). Les cours MAT 4774 et STA 4706 ne peuvent être combinés pour l'obtention de crédits. Antérieurement MAT 4774.

### **STA 4707 Méthodes de statistique multidimensionnelle (3 crédits)**

Vecteurs et matrices aléatoires. Loi normale multivariée et ses propriétés. Inférence dans le cadre de la loi normale multivariée: estimation, tests d'hypothèses. Analyse en composantes principales. Classification. Analyse discriminante. Corrélation canonique. Analyse des données à l'aide de logiciels de statistique.

**Volet :** Cours magistral

Préalables : (MAT 2541 ou MAT 2742), (MAT 3775 ou STA 3700). Les cours MAT 4775 et STA 4707 ne peuvent être combinés pour l'obtention de crédits. Antérieurement MAT 4775.

### **STA 4720 Chapitres choisis de statistique (3 crédits)**

Chapitres choisis en statistique.

**Volet :** Cours magistral

Préalable : MAT 2775 ou STA 2500 (des préalables supplémentaires peuvent s'appliquer selon le sujet du cours). Antérieurement MAT 4776.

### **STA 4740 Laboratoire de statistique (3 crédits)**

Ce cours offre aux étudiants l'occasion d'appliquer à des données réelles les techniques statistiques et la théorie apprises dans leur programme. Grâce à des études de cas, les étudiants proposeront et mettront en œuvre des solutions à des problèmes issus de la pratique. Les travaux de cours comprendront des discussions en classe, du travail en équipe, des méthodes de recherche et de lecture d'articles scientifiques, d'analyse exploratoire de données, de formulation de plans d'analyse statistique, d'analyse de données, de gestion de projet et les flux de travail reproductibles, d'interprétation des résultats de l'analyse statistique, de rédaction de rapports d'analyse statistique et de présentations orales.

Acquérir des compétences en consultation statistique. Réservé aux étudiants inscrits au B.Sc. spécialisé en statistique de 4e année. MAT 4783 et STA 4740 ne peuvent être combinés pour l'obtention de crédits.

**Volet :** Cours magistral

Préalables : (MAT 3575 ou STA 3500), (MAT 3775 ou STA 3700) et six crédits de cours parmi MAT 3778, MAT 3779, MAT 4774, MAT 4775, MAT 4778, MAT 4780, MAT 4782, STA 3701, STA 3702, STA 4706, STA 4707, STA 4703, STA 4702, STA 4704. Antérieurement MAT 4783.

### **STA 4900 Projet de recherche de premier cycle / Undergraduate Research Project (3 crédits / 3 units)**

L'étudiant(e) réalisera un projet de recherche d'un semestre sous la supervision d'un ou plusieurs membres du département. L'étudiant(e) écrira un rapport sur sa recherche et présentera un séminaire sur le sujet de son projet et sa contribution. Ce cours est conçu pour l'étudiant(e) intéressé(e) à poursuivre des études supérieures. / The student will undertake a one-semester project carried out under the supervision of one or more faculty members in the department. The student will write a report and present a seminar on some aspect of the research and his/her participation. This course is intended for students interested in pursuing graduate studies.

**Volet / Course Component:** Recherche / Research

Prerequisites: Student must be in the 4th year of an Honours program in Statistics, with a minimum CGPA of 7.0, and obtain the permission of the Department.