

---

# BIF7105

## Méthodes statistiques en bioinformatique

Plan de cours

### Responsable(s) du cours

---

**Coordination** : MAKARENKOV, Vladimir  
PK-4815  
[makarenkov.vladimir@uqam.ca](mailto:makarenkov.vladimir@uqam.ca)

#### Enseignement :

TAHIRI, Nadia  
PK-4660  
[tahiri.nadia@uqam.ca](mailto:tahiri.nadia@uqam.ca)

WILLEMS, Matthieu  
PK-4660  
[willems.matthieu@uqam.ca](mailto:willems.matthieu@uqam.ca)

### Description du cours

---

Fournir les bases statistiques pour énoncer les hypothèses statistiques et exécuter les tests qui s'appliquent à des problèmes de bioinformatiques.

Préalables académiques :

Rappel des principes de base en probabilité et en statistique. Inférence statistique classique et bayésienne. Processus stochastiques : promenades aléatoires, processus de Poisson, chaînes de Markov. Algorithme BLAST ("Basic Local Alignment Search Tool"). Modèles d'évolution.

### Objectif du cours

---

- 1) Appliquer et comprendre les bases statistiques en bioinformatique pour énoncer des hypothèses.
- 2) Apprendre le langage R et développer des fonctions d'analyse.

### Modalités d'évaluation

---

| ÉVALUATIONS          | DATES DE REMISE | PONDÉRATION |
|----------------------|-----------------|-------------|
| Examen de mi-session | 04 novembre     | 30%         |
| T.P. (devoir)        | 02 décembre     | 30%         |
| Projet de session    | 18 décembre     | 30%         |
| Présentation         | 16 décembre     | 10%         |

### TRAVAUX PRATIQUES (T.P.)

Le T.P. sera une suite de petits exercices qui touchent directement la matière enseignée les semaines précédentes. Les énoncés seront donnés trois semaines avant la date de remise, soit le 11 NOVEMBRE 2020. Les travaux pratiques doivent être faits individuellement et remis sous format papier.

### PROJET DE SESSION

Les projets de session sont à faire en ÉQUIPE DE DEUX OU TROIS ÉTUDIANTS. Les équipes devront définir et envoyer un énoncé de projet au plus tard le 11 novembre 2020. Chaque équipe devra rencontrer le professeur responsable de son projet au moins une fois durant la session pour discuter l'avancement de leur projet.

Le projet réalisé sera livré de deux façons : un exposé en classe de 10-15 MINUTES (5 minutes par personne) le 16 décembre 2020, puis un rapport écrit sous forme d'article à remettre le 18 décembre 2020. Vous devez remettre toutes les annexes que vous avez générées dans un dossier compressé (.zip ou .tar.gz). N'oubliez pas de bien les identifier et les commenter.

Le rapport principal sera de 10 pages maximum et comprendra les sections suivantes :

- 1) Un résumé de votre étude (1/2 page)
- 2) Une introduction avec les références pertinentes (2 pages)
- 3) Méthodologie (2 à 3 pages)
- 4) Résultats et discussions (2 à 3 pages)
- 5) Conclusion (1 page)
- 6) Références en écartant les références de *Wikipedia* (1 page maximum)

### EXEMPLE DE TRAVAUX ADMISSIBLES

- 1) Révision critique d'un outil bioinformatique (ou package R) : description de l'application et des jeux de données, de la problématique, des tests d'hypothèses et de la méthode statistique. Inclusion d'un exemple pratique.
- 2) Révision des techniques statistiques utilisées dans un sujet d'intérêt en bioinformatique p.ex. la phylogénie, la variation génétique, l'épidémiologie, l'écologie, etc.

## Calendrier

---

| Semaine | Date         | Activité   |
|---------|--------------|--|
| 1       | 9 septembre  | Présentation du cours, introduction aux statistiques, concept de probabilité, méthode scientifique, mesures de tendance centrale, mesures de dispersion et de position (MW). Présentation de l'environnement de programmation et introduction aux concepts algorithmiques (NT) |
| 2       | 16 septembre | Distributions de probabilité (loi de Poisson, loi normale, loi de Student), tests d'hypothèses, notion de degrés de liberté (MW)   |
| 3       | 23 septembre | Introduction au langage R avec RStudio, en R : Introduction aux structures de données : vecteurs, « data.frame » et matrices, lecture de fichiers, calculs simples et sorties graphiques (NT)  |

| Semaine | Date         | Activité   |
|---------|--------------|--|
| 4       | 30 septembre | Analyse de variance (ANOVA), test de normalité, analyse factorielle (ANOVA à plusieurs facteurs) (MW)                  |
| 5       | 7 octobre    | Langage R avec RStudio : Opérations sur les listes, structures de contrôle et fonctions, tests statistiques en R (NT)  |
| 6       | 14 octobre   | Régressions simples et multiples (MW)  |
| 7       | 21 octobre   | Langage R avec RStudio : ANOVA et régression dans R (NT)   |
| 8       | 28 octobre   | Chaînes de Markov et théorie de l'échantillonnage, théorème de Bayes et p-value (MW)                                   |
| 9       | 4 novembre   | Examen de mi-session (MW)  |
| 10      | 11 novembre  | Langage R avec RStudio : Analyse Bayésienne et analyse par composantes principales (PCoA) et distribution du T.P. (NT) |

---

| Semaine | Date        | Activité  |
|---------|-------------|---|
| 11      | 18 novembre | Simulations Monte Carlo, tests de permutation, bootstrap, concept de Likelihood (MW)  |
| 12      | 25 novembre | Langage R avec RStudio : Quantitative structure-activity relationship - QSAR (NT)   |
| 13      | 2 décembre  | Langage R avec RStudio : Simulations de Monte Carlo, échantillonnage aléatoire, bootstrap, réalisation de rapports, présentations avec Knitr et remise du T.P. (NT) |
| 14      | 9 décembre  | Langage R avec RStudio : Classification hiérarchique (hclust) et par partitionnement et k-means (NT)  |
| 15      | 16 décembre | Présentations orales (NT, MW)   |

---

## Médiagraphie

---

### LOGICIEL

RStudio : <https://www.rstudio.com/products/rstudio/download/> LO

### LIVRES ET AIDE MÉMOIRE

Kauffman, M. (2009) AIDE MÉMOIRE R Référence des fonctions de R les plus courantes. ([https://cran.r-project.org/doc/contrib/Kauffmann\\_aide\\_memoire\\_R.pdf](https://cran.r-project.org/doc/contrib/Kauffmann_aide_memoire_R.pdf)) VO

Krijnen, W. P. (2009). Applied statistics for bioinformatics using R. Institute for Life Science and Technology, Hanze University. (<https://cran.r-project.org/doc/contrib/Krijnen-IntroBioInfStatistics.pdf>) VO

Laude, H. (2016). Data Scientist et langage R - Guide d'autoformation à l'exploitation des Big Data. ENI. VR

Paradis, E. (2005) R pour les débutants ([https://cran.r-project.org/doc/contrib/Paradis-rdebuts\\_fr.pdf](https://cran.r-project.org/doc/contrib/Paradis-rdebuts_fr.pdf)) VO

### Participation à un cours ou à une activité d'enseignement en ligne

- Lors d'un cours ou d'une activité d'enseignement en ligne, le personnel enseignant peut décider, selon le cas, de procéder à l'enregistrement audio ou audiovisuel du cours ou de l'activité d'enseignement. Le personnel enseignant peut partager l'enregistrement uniquement à son groupe-cours.
- En cas d'enregistrement, l'étudiante, l'étudiant sera informé au début de la séance.
- Il est de la responsabilité de l'étudiante, de l'étudiant de désactiver son microphone et/ou sa caméra s'il ne souhaite pas être enregistré.
- À défaut de désactiver son microphone et/ou sa caméra, l'étudiante, l'étudiant consent à l'enregistrement audio ou audiovisuel, à la conservation, à la rediffusion et à l'utilisation de l'enregistrement de son nom, de sa voix et de son image dans le cadre du cours ou de l'activité en ligne. L'étudiante, l'étudiant reconnaît ne détenir aucun droit dans l'enregistrement.
- Sauf avec l'autorisation expresse écrite du personnel enseignant, il est interdit de reproduire, d'enregistrer, de publier, de diffuser, de communiquer ou de partager, par quelque moyen que ce soit, tout ou partie de l'enregistrement d'un cours ou d'une activité d'enseignement en ligne de même que tout matériel pédagogique s'y rattachant.
- Une étudiante, un étudiant qui contrevient à ce qui précède s'expose aux sanctions prévues dans les règlements et politiques de l'UQAM ou à tout recours légal, notamment en vertu de la Loi sur le droit d'auteur.

### Politique d'absence aux examens

*Reprise d'examen.* L'autorisation de reprendre un examen en cas d'absence est de **caractère exceptionnel**. Pour obtenir un tel privilège, l'étudiant.e doit avoir des motifs sérieux et bien justifiés.

*Conflits d'horaire.* Il est de la responsabilité de l'étudiant.e de ne pas s'inscrire à des cours qui sont en conflit d'horaire, tant en ce qui concerne les séances de cours ou d'exercices que les examens. **De tels conflits d'horaire ne constituent pas un motif justifiant une demande d'examen de reprise.**

*Procédure.* L'étudiant.e absent.e lors d'un examen doit, dans les cinq (5) jours ouvrables suivant la date de l'examen, présenter une demande de reprise en utilisant le formulaire prévu, disponible sur <http://info.uqam.ca/politiques>.

L'étudiant.e doit déposer le formulaire dûment complété au secrétariat de la direction de son programme d'études :

- PK-3150 pour les programmes de premier cycle
- PK-4150 pour les programmes de cycles supérieurs

*Pièces justificatives.* Dans le cas d'une absence pour raison médicale, l'étudiant.e doit joindre un certificat médical original et signé par le médecin décrivant la raison de l'absence à l'examen. Les dates d'invalidité doivent être clairement indiquées sur le certificat. Une vérification de la validité du certificat pourrait être faite. Dans le cas d'une absence pour une raison non médicale, l'étudiant.e doit fournir les documents originaux expliquant et justifiant l'absence à l'examen ; par exemple, lettre de la Cour en cas de participation à un jury, copie du certificat de décès en cas de décès d'un proche, etc. Toute demande incomplète sera refusée. Si la direction du programme d'études de l'étudiant.e constate qu'un.e étudiant.e a un comportement récurrent d'absence aux examens, l'étudiant.e peut se voir refuser une reprise d'examen.

*Pour plus d'informations.* Consulter la page <http://info.uqam.ca/politiques>.



### **Règlement numéro 18 sur les infractions de nature académique (extraits)**

Tout acte de plagiat, fraude, copiage, tricherie ou falsification de document commis par une étudiante, un étudiant, de même que toute participation à ces actes ou tentative de les commettre, à l'occasion d'un examen ou d'un travail faisant l'objet d'une évaluation ou dans toute autre circonstance, constituent une infraction au sens de ce règlement.

La liste non limitative des infractions est définie comme suit :

- la substitution de personnes ;
- l'utilisation totale ou partielle du texte d'autrui en la faisant passer pour sien ou sans indication de référence ;
- la transmission d'un travail pour fins d'évaluation alors qu'il constitue essentiellement un travail qui a déjà été transmis pour fins d'évaluation académique à l'Université ou dans une autre institution d'enseignement, sauf avec l'accord préalable de l'enseignante, l'enseignant ;
- l'obtention par vol, manoeuvre ou corruption de questions ou de réponses d'examen ou de tout autre document ou matériel non autorisés, ou encore d'une évaluation non méritée ;
- la possession ou l'utilisation, avant ou pendant un examen, de tout document non autorisé ;
- l'utilisation pendant un examen de la copie d'examen d'une autre personne ;
- l'obtention de toute aide non autorisée, qu'elle soit collective ou individuelle ;
- la falsification d'un document, notamment d'un document transmis par l'Université ou d'un document de l'Université transmis ou non à une tierce personne, quelles que soient les circonstances ;
- la falsification de données de recherche dans un travail, notamment une thèse, un mémoire, un mémoire-crédation, un rapport de stage ou un rapport de recherche ;
- Les sanctions reliées à ces infractions sont précisées à l'article 3 du Règlement no 18.

Les règlements concernant le plagiat seront strictement appliqués. Pour plus de renseignements :

- <http://www.infosphere.uqam.ca/rediger-un-travail/eviter-plagiat>
- <http://r18.uqam.ca/>

### **Politique no 16 visant à prévenir et combattre le sexisme et les violences à caractère sexuel**

Les violences à caractère sexuel se définissent comme étant des comportements, propos et attitudes à caractère sexuel non consentis ou non désirés, avec ou sans contact physique, incluant ceux exercés ou exprimés par un moyen technologique, tels les médias sociaux ou autres médias numériques. Les violences à caractère sexuel peuvent se manifester par un geste unique ou s'inscrire dans un continuum de manifestations et peuvent comprendre la manipulation, l'intimidation, le chantage, la menace implicite ou explicite, la contrainte ou l'usage de force.

Les violences à caractère sexuel incluent, notamment :

- la production ou la diffusion d'images ou de vidéos sexuelles explicites et dégradantes, sans motif pédagogique, de recherche, de création ou d'autres fins publiques légitimes ;
- les avances verbales ou propositions insistantes à caractère sexuel non désirées ;
- la manifestation abusive et non désirée d'intérêt amoureux ou sexuel ;
- les commentaires, les allusions, les plaisanteries, les interpellations ou les insultes à caractère sexuel, devant ou en l'absence de la personne visée ;
- les actes de voyeurisme ou d'exhibitionnisme ;
- le (cyber) harcèlement sexuel ;
- la production, la possession ou la diffusion d'images ou de vidéos sexuelles d'une personne sans son consentement ;
- les avances non verbales, telles que les avances physiques, les attouchements, les frôlements, les pincements, les baisers non désirés ;
- l'agression sexuelle ou la menace d'agression sexuelle ;
- l'imposition d'une intimité sexuelle non voulue ;
- les promesses de récompense ou les menaces de représailles, implicites ou explicites, liées à la satisfaction ou à la non-satisfaction d'une demande à caractère sexuel.

*Pour consulter la politique no 16*

[https://instances.uqam.ca/wp-content/uploads/sites/47/2018/05/Politique\\_no\\_16.pdf](https://instances.uqam.ca/wp-content/uploads/sites/47/2018/05/Politique_no_16.pdf)

*Pour obtenir de l'aide, faire une divulgation ou une plainte*

Bureau d'intervention et de prévention en matière de harcèlement  
514-987-3000, poste 0886

*Pour obtenir la liste des services offerts à l'UQAM et à l'extérieur de l'UQAM*

<https://harcelement.uqam.ca>

*Soutien psychologique (Services à la vie étudiante)*

514-987-3185  
Local DS-2110

*CALACS Trêve pour Elles – point de services UQAM*

514 987-0348  
[calacs@uqam.ca](mailto:calacs@uqam.ca)  
<http://trevepourelles.org>

*Service de la prévention et de la sécurité*

514-987-3131

### **Politique no 44 d'accueil et de soutien des étudiant.e.s en situation de handicap**

*Politique.* Par sa politique, l'Université reconnaît, en toute égalité des chances, sans discrimination ni privilège, aux étudiant.e.s en situation de handicap, le droit de bénéficier de l'ensemble des ressources du campus et de la communauté universitaire, afin d'assurer la réussite de leurs projets d'études, et ce, dans les meilleures conditions possibles. L'exercice de ce droit est, par ailleurs, tributaire du cadre réglementaire régissant l'ensemble des activités de l'Université.

*Responsabilité de l'étudiant.e.* Il incombe aux étudiant.e.s en situation de handicap de rencontrer les intervenant.e.s (conseiller.ère.s à l'accueil et à l'intégration du Service d'accueil et de soutien des étudiant.e.s en situation de handicap, professeur.e.s, chargé.e.s de cours, direction de programmes, associations étudiantes concernées, etc.) qui pourront faciliter leur intégration à la communauté universitaire ou les assister et les soutenir dans la résolution de problèmes particuliers en lien avec les limitations entraînées par leur déficience.

*Service d'accueil et de soutien aux étudiant.e.s en situation de handicap.* Le Service d'accueil et de soutien aux étudiant.e.s en situation de handicap (SASESH) offre des mesures d'aménagement dont peuvent bénéficier certains étudiant.e.s. Il est fortement recommandé aux de se prévaloir de ces services afin de réussir ses études, sans discrimination. Pour plus d'information, visiter le site de ce service : <https://vie-etudiante.uqam.ca/etudiant-situation-handicap/nouvelles-ressources.html> et celui de la politique institutionnelle d'accueil et de soutien aux étudiant.e.s en situation de handicap : [https://instances.uqam.ca/wp-content/uploads/sites/47/2018/05/Politique\\_no\\_44.pdf](https://instances.uqam.ca/wp-content/uploads/sites/47/2018/05/Politique_no_44.pdf)

Il est important d'informer le SASESH de votre situation le plus tôt possible :

- En personne : 1290, rue Saint-Denis, Pavillon Saint-Denis, local AB-2300
- Par téléphone : 514 987-3148
- Par courriel : [situation.handicap@uqam.ca](mailto:situation.handicap@uqam.ca)
- En ligne : <https://vie-etudiante.uqam.ca/>