

INFORMATIONS SUR LE COURS							
Code et titre du cours :	201-424-SL – Biostatistique	Durée :	60 h	Pondération :	2-2-2	Unités :	2,00
Préalable(s) :	SN ou TS secondaire 5 Il est préférable de faire <i>Géomatique et traitement de données</i> (145-223-SL) avant.	Session :	4	Discipline :	Mathématiques		
PRÉSENTATION GÉNÉRALE DU COURS							
Contribution au programme :	<p>Les compétences et les notions acquises dans ce cours seront appliquées dans les cours d’initiation à la recherche.</p> <p>La compétence 043W est partagée avec les cours Initiation à la recherche 1 (145-581-SL) et Initiation à la recherche 2 (145-682-SL).</p> <p>La compétence 044A est partagée avec les cours Géomatique et traitement de données (145-223-SL), Initiation à la recherche 1 (145-581-SL) et Initiation à la recherche 2 (145-682-SL).</p>						
Description du cours :	<p>Dans ce cours, l’étudiant acquerra les outils statistiques nécessaires à l’établissement du portrait d’une situation. Les notions suivantes seront abordées : vocabulaire statistique, méthodes d’échantillonnage, tableaux, graphiques et mesures en lien avec des variables statistiques, estimation par intervalle de confiance, tests d’hypothèses, corrélation entre deux variables, valeurs extrêmes.</p>						
Objectif terminal du cours :	<p>Comprendre le vocabulaire statistique et appliquer certaines méthodes de travail et d’analyse statistique afin d’étudier les variations biologiques, les phénomènes qui s’y rattachent et les problèmes qui en découlent.</p>						
Lien avec le profil de sortie :	<p>Les compétences associées à ce cours font partie de l’axe de formation « Méthodologie » donc qui permet de développer des bonnes méthodes de travail en lien avec le traitement de données.</p>						
COMPÉTENCES VISÉES							
Code(s) et énoncé(s) de compétence(s) :	Éléments de compétence :					Atteinte complète ou partielle :	
043W – Établir le portrait d’une situation à l’aide de données statistiques	<ol style="list-style-type: none"> Effectuer la saisie informatisée des données. Vérifier la validité des données fournies. Déterminer la nature du traitement statistique nécessaire. Effectuer les traitements statistiques prévus. Assurer la mise en forme et la présentation graphique des données. Analyser et interpréter les résultats obtenus. 					Partielle	
044A – Utiliser les technologies numériques et informatiques dans le cadre de ses activités professionnelles	<ol style="list-style-type: none"> Gérer son environnement informatique. Exploiter des applications de traitement de textes. Utiliser des outils informatisés de traitement statistique et de représentation de données. Utiliser des outils de géomatique et des cartes géoréférencées. Utiliser différents appareils numériques à des fins d’exploitation d’images. Utiliser l’Internet et l’intranet à des fins de recherche et de communication. 					Partielle	

COMPOSITION DU COURS		
Énoncé de compétence :	Contexte de réalisation :	
043W – Établir le portrait d'une situation à l'aide de données statistiques	<ul style="list-style-type: none"> ✓ À partir de situations appliquées au domaine de la biologie. ✓ À l'aide de tables, de graphiques et de manuels de référence. ✓ À l'aide de logiciels appropriés, tel un chiffrier électronique. 	
Activités d'enseignement/ apprentissage	Théorie sous forme magistrale et exercices.	
Éléments de compétence	Critères de performance	Objectifs d'apprentissage et contenus essentiels
1. Effectuer la saisie informatisée des données.	1. Formulation de requêtes adéquates en fonction des bases de données.	1.1 Utiliser des fonctions de base d'un chiffrier (révision de notions vues dans le cours Géomatique et traitement de données (145-223-SL)) : <ul style="list-style-type: none"> - addition, soustraction, division, multiplication, logarithme, puissance; - moyenne, écart-type, rang, racine carrée; - calcul du nombre de cellules contenant des valeurs numériques. 1.2 Utiliser des fonctions de base d'un chiffrier (nouveau contenu): <ul style="list-style-type: none"> - générer un nombre aléatoire; - rechercher selon un critère; - utiliser un filtre; - construire des tableaux de distribution; - construire des tableaux croisés dynamiques; - construire des graphiques (par exemple, histogramme, diagramme à bâtons, courbe de fréquences cumulées).
	2. Utilisation adéquate des applications appropriées, incluant le formatage du chiffrier.	(voir ci-haut, critère de performance 1)
	3. Saisie précise et complète des données.	3.1 Saisir et présenter les données adéquatement.
2. Vérifier la validité des données.	1. Repérage des biais liés à la collecte des données.	1.1 Identifier les différentes sources de biais : <ul style="list-style-type: none"> - méthodes d'échantillonnage inadéquates; - erreurs de mesure; - instruments de mesure imprécis.
	2. Évaluation juste du degré de précision et de pertinence des mesures fournies.	2.1 Distinguer une donnée exacte d'une donnée précise. 2.2 Prendre en considération l'approximation des données.
	3. Rejet des données non valides et non pertinentes pour la poursuite des travaux.	3.1 Repérer les données expérimentales extrêmes parmi l'ensemble des données d'une expérience. 3.2 Déterminer et décrire le test à appliquer pour évaluer si une donnée (réplicat) est extrême : le Q-test (entre 3 et 30 données) ou méthode de l'étendue interquartile (si $n > 30$). 3.3 Appliquer le Q-test et la méthode de l'étendue interquartile pour évaluer si une donnée est extrême. 3.4 Discuter, en tenant compte du contexte de l'expérience, de si une donnée extrême peut être considérée comme étant aberrante et peut alors être rejetée d'une série de données (réplicats).

<p>3. Déterminer la nature du traitement statistique nécessaire.</p>	<p>1. Caractérisation correcte des échantillons : types d'échantillons, types de paramètres et types de distributions.</p>	<p>1.1 Identifier unité statistique, population, échantillon et variable (dépendante et indépendante).</p> <p>1.2 Identifier le type de données : qualitative (nominale ou ordinale), quantitative (discrète ou continue).</p> <p>1.3 Identifier et appliquer les différentes méthodes d'échantillonnage : aléatoire simple, systématique, stratifié et par grappes.</p> <p>Mesures de tendance centrale et de dispersion</p> <p>1.4 Identifier et différencier les mesures de tendance centrale : moyenne arithmétique, médiane et mode.</p> <p>1.5 Calculer et interpréter les mesures de tendance centrale.</p> <p>1.6 Identifier et différencier les mesures de dispersion : étendue, étendue interquartile, variance (d'une population = σ^2 et d'un échantillon = S^2), coefficient de variation et écart-type.</p> <p>1.7 Déterminer les quartiles Q_1 et Q_3 dans une série de données.</p> <p>1.8 Calculer et interpréter les mesures de dispersion.</p> <p>1.9 Nommer les avantages et les inconvénients des mesures de tendance centrale et des mesures de dispersion.</p> <p>1.10 Définir et appliquer les notions de base des probabilités : expérience, espace échantillonnal, événement, définition classique d'une probabilité.</p> <p>1.11 Calculer le degré de liberté (dl).</p> <p>1.12 Décrire les distributions d'échantillonnage :</p> <ul style="list-style-type: none"> - loi de Student (pour un échantillon plus petit que 30 ou σ est inconnu); - loi normale (pour un échantillon plus grand que 30 ou σ est connu); - loi binomiale (brièvement); - loi du Khi-carré. <p>1.13 Calculer et interpréter les intervalles de confiance : estimation de la moyenne et d'une proportion (ou fréquence relative).</p>
	<p>2. Sélection des traitements et des tests statistiques en fonction de la nature des données et de l'hypothèse de départ.</p>	<p>2.1 Décrire l'utilité et les conditions d'applications de :</p> <ul style="list-style-type: none"> - la comparaison de moyennes (test t, analyse de variance ou Anova); - la comparaison de proportions (Khi-carré – χ^2); - la corrélation et la régression linéaire simple. <p>2.2 Décrire le coefficient de corrélation (r) et de détermination (r^2).</p> <p>2.3 Appliquer un test d'hypothèse : définir hypothèse nulle (H_0), hypothèse(s) alternative(s) (H_1), type d'erreur (α et β), règle de décision.</p> <p>2.4 À partir d'une situation concrète liée à la biologie, déterminer les hypothèses, les erreurs α et β ainsi que le test à effectuer.</p>
<p>4. Effectuer les traitements statistiques prévus.</p>	<p>1. Sélection des fonctions statistiques appropriées.</p>	<p>(voir élément de compétence 3, critère de performance 2)</p>
	<p>2. Sélection des variables pertinentes.</p>	<p>2.1 Sélectionner les variables pertinentes dans un tableau de données selon le caractère étudié.</p>

	3. Exécution adéquate des fonctions sélectionnées et exactitude des calculs.	3.1 Appliquer les tests d'hypothèse suivants : <ul style="list-style-type: none"> - test sur une seule moyenne, test sur deux moyennes, test sur x moyennes (comparaison de moyennes : test t et Anova); - test sur une seule proportion, test sur deux proportions, test sur x proportions (comparaison de fréquences : Khi-carré – χ^2); - corrélation et régression linéaire simple. 3.2 Déterminer le coefficient de corrélation (r) et de détermination (r^2).
5. Assurer la mise en forme et la présentation graphique des données.	1. Sélection du mode de présentation graphique le plus approprié.	1.1 Différencier les types de graphiques : <ul style="list-style-type: none"> - diagramme en bâtons; - diagramme à bandes; - histogramme; - polygone de fréquences; - nuages de points. 1.2 Choisir le type de graphique le plus approprié à une situation donnée.
	2. Qualité des tableaux : <ul style="list-style-type: none"> - pertinence des données présentes; - prise en compte des paramètres appropriés (structure, organisation et lisibilité facilitant l'interprétation). 	2.1 Respecter les normes du manuel <i>Prise de données et rédaction en sciences biologiques</i> établi par le programme de bioécologie (vu dans le cours Méthodologie en bioécologie (145-121-SL)) pour la présentation des tableaux.
	3. Qualité des figures : <ul style="list-style-type: none"> - présentation des points saillants; - conformité aux normes de présentation. 	3.1 Respecter les normes du manuel <i>Prise de données et rédaction en sciences biologiques</i> établi par le programme de bioécologie (vu dans le cours Méthodologie en bioécologie (145-121-SL)) pour la présentation des figures.
6. Analyser et interpréter les résultats.	1. Définition correcte des différentes notions statistiques et de la terminologie utilisée.	(voir éléments de compétence 3 et 4)
	2. Justesse des explications relatives aux traitements statistiques effectués.	(voir éléments de compétence 3 et 4)
	3. Degré adéquat de précision et de signification des résultats présentés.	(voir éléments de compétence 3 et 4)
	4. Justesse de l'interprétation compte tenu de la problématique de départ.	(voir éléments de compétence 3 et 4)
Énoncé de compétence :		Contexte de réalisation :
044A – Utiliser les technologies numériques et informatiques dans le cadre de ses activités professionnelles		<ul style="list-style-type: none"> ✓ À partir de situations appliquées au domaine de la biologie. ✓ À l'aide de tables, de graphiques et de manuels de référence. ✓ À l'aide de logiciels appropriés, tel un chiffrier électronique.
Activités d'enseignement/ apprentissage	Théorie sous forme magistrale et exercices.	
Éléments de compétence	Critères de performance	Objectifs d'apprentissage et contenus essentiels

3. Utiliser des outils informatisés de traitement statistique et de représentation de données.	4. Saisie adéquate des données fournies.	4.1 Compiler des données à l'aide d'un chiffrier.
	5. Utilisation des fonctions de base relatives au traitement statistique de l'application.	5.1 Utiliser un chiffrier et un logiciel d'analyse de données pour toutes les fonctions statistiques vues dans ce cours.

ÉVALUATION TERMINALE DE COURS

Nature de l'évaluation terminale du cours :	Contexte de réalisation	Critères d'évaluation propres à l'évaluation terminale	Pondération
Évaluation où l'étudiant devra démontrer l'acquisition des outils statistiques pour dresser le portrait de diverses situations.	L'évaluation terminale de cours est individuelle, de type synthèse et doit inclure un examen. À des fins d'équité, au cours de chaque session, les professeurs qui se sont vus attribuer un même numéro de cours se concertent de manière à rendre l'évaluation équivalente.	<ul style="list-style-type: none"> - la qualité du déploiement d'un raisonnement mathématique - l'expression claire d'une démarche - le respect de la syntaxe de l'écriture mathématique - la rigueur dans la justification des étapes - l'exactitude des calculs 	Entre 25% et 40% de la note finale.

STRATÉGIES D'ÉVALUATION SUGGÉRÉES

Au moins 3 examens théoriques (environ 75%)
Au moins 2 laboratoires (environ 25%)

EXIGENCES PARTICULIÈRES (PIEA ET PDEA, ETC.)

Les exemples et les exercices doivent être appliqués à la biologie.
Respecter les normes de *Prise de données et rédaction en sciences biologiques* établi par le programme de bioécologie (vu dans le cours Méthodologie en bioécologie (145-121-SL)).
Respecter la Politique départementale d'évaluation des apprentissages (PDÉA) du département de mathématiques.

MÉDIAGRAPHIE :

Médiagraphie indicative, logiciels et/ou équipements suggérés :

Livres de référence

Colin, Michèle et Gisèle Payette (2004). *Biostatistiques pour les techniques biologiques*. Publié par les auteurs. Sherbrooke.
Grenon, Gilles et Suzanne Viau (1997). *Statistiques appliquées : initiation à l'analyse des données statistiques*. Gaëtan Morin éditeur.
Howard, Christensen (1986). *La statistique : démarche pédagogique programmée*. Gaëtan Morin éditeur.
Laliberté, Célyne (2005). *Probabilité et statistiques : de la conception à la compréhension*. ERPI.
Mercier, André (1988). *Éléments de biométrie*. Les éditions Le Griffon d'argile.
Ouellet, Gilles (1998). *Statistique et probabilité*. Les éditions Le Griffon d'argile.
Scherrer, Bruno (2007). *Biostatistique 2^e édition*. Montréal : Chenelière éducation, 832 p.
Simard, Christiane (2002). *Notions de statistiques*. Les éditions Le Griffon d'argile.

Bastien, Claudia; Fréchette, Jean-Martin; Gauvin, Marie-Josée (2017). *Prise de données et rédaction en sciences biologiques*. Montréal : Chenelière éducation, 240 p.

Logiciels

Excel, SigmaPlot, Logiciel R, Statview, SAS, Systat, JMP, SPSS