

SOL1020- Automne 2022

INTRODUCTION À LA STATISTIQUE SOCIALE

© El Hadj Touré, Ph.D. Sociologie

Horaire du cours : Théorie : Vendredi, 13h-15h, B-3315 Pav. 3200 J. Brillant
Labo : Vendredi, 15h-17h, B-1215 Pav. 3200 J-Brillant

Disponibilité du professeur : Sur rendez-vous

Courrier électronique : elhadj.toure@umontreal.ca

Auxiliaire : yao.robert.djogbenou@umontreal.ca

Espace virtuel : <https://studium.umontreal.ca/user/index.php?id=228214>

DESCRIPTION DU COURS

Processus de développement d'une recherche quantitative. Concepts de base en statistiques, fréquences et pourcentages, mesures de tendance centrale et de dispersion, tableaux croisés. Introduction à l'inférence statistique, test chi-carré, test t, ANOVA, corrélation et régression. Initiation au logiciel d'analyse statistique SPSS et utilisation du tableur Excel.

OBJECTIFS D'APPRENTISSAGE

Le cours d'*introduction à la statistique sociale* présente les *méthodes, techniques* et *procédures* statistiques de base appliquées aux sciences sociales, de façon à amener les étudiant-e-s à pouvoir répondre à des questions de recherche descriptives et relationnelles. À la fin du cours, les étudiants-e-s seront en mesure de :

- 1) Connaître les étapes sous-jacentes au cheminement d'une recherche quantitative et savoir situer la place qu'y occupent les statistiques ;
- 2) Décrire un ensemble de données se rapportant à une ou deux variables à l'aide de statistiques, tableaux et graphiques appropriés ;
- 3) Procéder à l'inférence statistique en utilisant de façon adéquate les techniques d'estimation et de validation d'hypothèses les plus courantes en sciences sociales ;
- 4) Effectuer ces analyses univariées et bivariées à l'aide des logiciels Excel et SPSS et être en mesure d'en interpréter statistiquement et théoriquement les résultats ;

- 5) Savoir émettre une critique méthodique de travaux de recherche faisant appel aux statistiques descriptives et inférentielles.

DÉMARCHE PÉDAGOGIQUE

Dans le but d'atteindre les objectifs susmentionnés, chaque séance hebdomadaire est divisée en deux temps. La première période consiste en un *exposé magistral* entrecoupé de remue-méninges et d'exercices « éclair ». Sans tomber dans le piège du formalisme mathématique outrancier, le professeur présente les statistiques sociales de base selon une approche intuitive. Derrière chaque formule, en effet, se cache une intelligence logico-mathématique et conceptuelle. On peut y accéder à l'aide de mises en situation, de simulations simples et d'illustrations visuelles. Ce n'est toutefois pas un cours de maths. Les maths sont plutôt utilisées pour donner sens aux phénomènes sociaux.

La deuxième période, quant à elle, est consacrée à l'application en laboratoire informatique des connaissances théoriques préalablement acquises. À cet effet, les étudiant-e-s utiliseront le tableur Excel pour effectuer des opérations de calcul. Nous sommes résolument d'avis que pour comprendre pourquoi les statistiques font ce qu'elles font, il faut certes savoir les calculer, avant de pouvoir les interpréter convenablement dans le contexte des sciences sociales. Les étudiant-e-s feront surtout l'apprentissage du logiciel SPSS afin de procéder automatiquement à des analyses statistiques uni-bivariées. Ils pourront ainsi se familiariser avec cet outil informatique en manipulant des données sociales réelles. Pour effectuer les travaux pratiques à la maison, les étudiant-e-s peuvent installer SPSS sur leur ordinateur personnel.¹

Somme toute, la pédagogie est axée sur la compréhension conceptuelle et l'interprétation contextualisée des statistiques sociales de base. Même si les statistiques se présentent comme des outils sophistiqués à utiliser avec « imagination sociologique » (Mills, 1959), elles peuvent être instrumentalisées à volonté, comme nous le verrons. Afin de faciliter leur apprentissage, les étudiant-e-s devront se préparer par les lectures hebdomadaires suggérées dans le volume de référence. De plus, après chaque séance hebdomadaire, ils pourront s'autoévaluer en ligne à travers un *test quiz* et réaliser une série de *dix exercices récapitulatifs corrigés*. Les applications avec les logiciels sont illustrées visuellement dans un cahier de laboratoire, de telle sorte que les étudiant-e-s peuvent suivre facilement les démonstrations du professeur et réaliser les exercices pratiques proposés.

SOPPORTS DE COURS

Avant chaque séance hebdomadaire, les étudiant-e-s ont la possibilité de télécharger les présentations PowerPoint (sous divers formats) et les applications à effectuer au labo (Excel et SPSS). Les applications logicielles sont illustrées visuellement dans un cahier de laboratoire, de telle sorte que les étudiants peuvent suivre facilement les démonstrations du professeur et réaliser les exercices proposés. Tous ces supports de cours sont disponibles sur StudiUM.

Au demeurant, les notes de cours résument certes l'essentiel de la partie théorique du cours. Mais, elles n'en fournissent pas le contenu intégral, le style télégraphique étant utilisé à bien des égards dans les présentations PowerPoint. Par conséquent, la présence en classe et au labo est obligatoire pour maîtriser la matière et augmenter les chances de réussir le cours.

¹ https://studium.umontreal.ca/pluginfile.php/5530230/mod_resource/content/30/SPSS-installation.pdf

L'achat du volume de référence (Fox, 1999) est obligatoire. Les étudiants peuvent acheter leur livre sur place à la librairie Jean-Brillant ou passer leur commande sur le site de la librairie. Ci-dessous un lien qui explique comment obtenir vos livres : <https://www.librairie.umontreal.ca/default.aspx>.

MODALITÉS D'ÉVALUATION

Réalisés individuellement ou tout au plus en équipe de trois, deux *travaux pratiques* (TPs) seront proposés aux étudiant-e-s tout au long de ce cours. Les TPs consistent en des analyses à effectuer à l'aide du logiciel SPSS afin de répondre à des questions de recherche. Ils seront remis en format « numérique » sur Studium dans l'espace prévu à cet effet, conformément à la programmation.

Un *examen-intra*, qui portera sur les statistiques descriptives, et un *examen final*, qui portera sur les statistiques inférentielles, seront organisés sur table. Ils se feront à « livres et cahiers ouverts ». Les examens sont individuels : ils consistent en des questions de calcul et problèmes pratiques à résoudre ainsi qu'en l'interprétation de résultats statistiques.

Les quiz 2 à 11 sont **obligatoires** et les étudiant-e-s les réaliseront directement sur Studium, tout au plus une semaine après leur disponibilité. Chacun des quiz porte sur 20 questions objectives de compréhension en lien avec une leçon. Les questions se présentent sous diverses formes : vrai/faux, choix multiples, numérique, appariement, etc. Le quiz 1 est **optionnel**.

Ci-dessous un récapitulatif des composantes de l'évaluation et leur échéancier :

Items	Date de disponibilité	Date de remise	Pondération
10 Quiz (2 à 11)	Après chaque leçon	Une semaine après	20%
TP1	23 septembre (17h)	12 octobre (23h59)	15%
Examen intra	21 octobre (13h-16h)		25%
TP2	11 novembre (17h)	2 décembre (23h59)	15%
Examen final	16 décembre (13h-16h)		25%
Total	--	--	100%

NB : La calculatrice scientifique est indispensable à la réalisation des travaux, tout comme à la réalisation des exercices « éclair » proposés dans le cours théorique.

Remise des travaux

Aucun retard dans la remise des travaux ne sera toléré, à moins d'une raison sérieuse acceptée par le professeur. *Les retards non autorisés seront pénalisés.* Une pénalité de 10 % de la note du travail en question sera appliquée par jour de retard.

Qualité de la langue

La notation prend en compte autant la justesse des calculs que la qualité de la rédaction dans les réponses et interprétations. Lorsque la qualité de la langue française fait défaut, une pénalité s'applique jusqu'à concurrence de 10% de la note obtenue.

Méfiez-vous du plagiat

Le plagiat, sous toutes ses formes², est pris très au sérieux par le professeur. Il ne sera aucunement toléré lorsqu'il dépasse le seuil minimal de 20%.

² [Questions_Quiz_V4-2013-04-question_avec_reponses.pdf \(usherbrooke.ca\)](https://www.usherbrooke.ca/Questions_Quiz_V4-2013-04-question_avec_reponses.pdf)





Dates limites





21 septembre : Modification d'un choix de cours et abandon d'un cours sans frais





11 novembre: Abandon d'un cours avec frais





DÉROULEMENT DU COURS




Sous réserve de modification au cours du semestre, le contenu et le calendrier des activités du cours sont ainsi définis :




Semaine 1	
Vend. 09 sept.	<ul style="list-style-type: none"> • L'enseignement et l'apprentissage des statistiques en sciences sociales • Présentation du cours et des modalités d'évaluation <p> Leçon 1. Cheminement d'une recherche quantitative</p> <ul style="list-style-type: none"> • Place des statistiques en recherche quantitative • Construction d'une problématique sous l'angle de l'approche quantitative • Opérationnalisation de l'hypothèse : des concepts aux indicateurs • Aperçu des techniques de collecte et d'analyse statistique des données • Interprétation statistique et sociologique des résultats <p><i>Lectures suggérées : Fox, pp.24-28 ; Lamoureux (2003)</i></p> <p>Quiz 1 à faire (optionnel, à titre expérimental)</p>
	<p> Labo 1. Introduction à l'utilisation du site web du cours (accès, planification et contenu du cours, test quiz, glossaire, évaluation). Initiation au tableur Excel : réaliser quelques opérations de calcul de base, manipuler les cellules de calcul, comprendre ce que sont les références absolues et relatives. Administration d'un sondage web auprès des étudiants en vue de la constitution d'une base de données SPSS.</p>
Semaine 2	
Vend. 16 sept.	<p> Leçon 2. Statistiques et variables</p> <ul style="list-style-type: none"> • Statistique vs statistiques, population vs échantillon, paramètre vs statistique, statistiques descriptives vs statistiques inférentielles • Qu'est-ce que les statistiques sociales ? Lien historique entre statistiques et sociologie • Classification des variables selon le type (quantitatif vs qualitatif) et le niveau de mesure (nominal, ordinal, intervallaire, proportionnel) • Des catégories mutuellement exclusives et collectivement exhaustives • Préalables à l'utilisation des statistiques : codage et quantification <p><i>Lectures suggérées : Fox, pp. 4-31</i></p> <p>Quiz 2 à faire (date limite 23 septembre, 23h59)</p>
	<p> Labo 2. Utilisation du tableur Excel pour créer et interpréter une table d'indice de masse corporelle, représenter visuellement les échelles de mesure. Introduction au</p>

	logiciel statistique SPSS : créer une base de données, afficher les variables et l'éditeur de données, manipuler et sauvegarder les données, vérifier les données et variables, procéder à quelques traitements liminaires, sauvegarder les résultats, changer certaines options.
Semaine 3	
Vend. 23 sept.	 Leçon 3. Distributions de fréquences et de pourcentages <ul style="list-style-type: none"> • Situations de recherche décrivant l'ampleur d'un phénomène social • Transformation des fréquences simples en fréquences regroupées : l'application de la formule de Sturges • Conversion des fréquences en %, puis les % en % cumulés • Mesures de rapport : proportion, %, taux, % de variation, ratio, indice • Construction de tableaux et de graphiques appropriés, et interprétation <p><i>Lectures suggérées : Fox, pp. 33-68</i></p> <p>Quiz 3 à faire (date limite 30 septembre, 23h59)</p>
	 Labo 3. À l'aide des logiciels Excel et SPSS, dépouiller les données d'une variable selon une distribution de fréquences et de %, les résumer au moyen d'un tableau et de graphiques appropriés (secteurs, diagramme en bâtons, diagramme en barres verticales et horizontales, histogramme, polygone, ogive). À l'aide de SPSS, « recoder » des variables qualitative et quantitative de façon à réduire leurs valeurs, sortir leur distribution, et traiter les valeurs manquantes. Lire et interpréter les sorties SPSS, présenter les résultats sous forme textuelle et dans des tableaux publiables à l'aide du logiciel Word. <p>TP1 à faire : Statistiques et variables, fréquences et pourcentages, mesures de tendance centrale et de variation</p>
Semaine 4	
Vend. 30 sept.	 Leçon 4. Mesures de tendance centrale <ul style="list-style-type: none"> • Situations de recherche décrivant la représentativité ou le caractère typique d'un phénomène social • Calcul et interprétation des mesures de tendance centrale les plus rapportées : mode, médiane et moyenne • Effet des cas déviants sur la tendance centrale • Types de distribution : unimodale, bimodale, plurimodale, symétrique, asymétrique positivement, asymétrique négativement • Quelle mesure de tendance centrale choisir ? Mise en situation pratique <p><i>Lectures suggérées : Fox, pp. 69-90</i></p> <p>Quiz 4 à faire (date limite 07 octobre, 23h59)</p>
	 Labo 4. À l'aide d'Excel, exécuter des calculs concernant le mode, la médiane et la moyenne. À l'aide de SPSS, effectuer ces opérations de façon automatique, représenter les mesures de tendance centrale au moyen d'un diagramme en boîte, détecter et exclure les cas déviants, « sélectionner » des observations, « calculer une variable » à partir

	d'autres variables. Lire et interpréter les sorties SPSS, présenter les résultats sous forme textuelle et dans des tableaux publiables à l'aide du logiciel Word.
Semaine 5	
Vend. 07 oct.	 Leçon 5. Mesures de variation <ul style="list-style-type: none"> • Situations de recherche décrivant la variabilité d'un phénomène social • Calcul et interprétation des mesures de variation les plus rapportées : étendue, intervalle interquartile, variance, écart-type et coefficient de variation • Mesure de la forme d'une distribution : symétrie et kurtose • Les scores-z ou scores standardisés et leur utilité pratique et statistique • La loi normale et ses applications pratiques <p><i>Lectures suggérées : Fox, pp. 91-107</i></p> <p>Quiz 5 à faire (date limite 14 octobre, 23h59)</p>
	 Labo 5. À l'aide d'Excel, calculer les mesures de variation, construire une table statistique de la distribution normale et appliquer la loi normale à des situations pratiques. À l'aide de SPSS, sortir les mesures de variation de façon automatique, les représenter au moyen d'un histogramme (avec courbe gaussienne) et d'un diagramme en boîte et à moustaches (pour détecter les cas déviants), « scinder » un fichier afin de comparer des groupes selon la dispersion d'une variable quantitative. Lire et interpréter les sorties SPSS en termes de variabilité, présenter les résultats sous forme textuelle et dans des tableaux publiables à l'aide du logiciel Word.
Semaine 6	
Vend. 14 oct.	 Leçon 6. Analyse de tableaux bivariés <ul style="list-style-type: none"> • Situations de recherche et conditions d'application de l'analyse de tableaux croisés • Décrire des données mettant en relation deux variables qualitatives à l'aide d'un tableau croisé en fréquences, puis calculer les pourcentages • Déterminer l'existence, l'intensité, la direction et la forme d'une relation • Analyse des diagrammes en barres groupées et empilées • L'association bivariée n'implique pas un lien de causalité <p><i>Lectures suggérées : Fox, pp. 121-147.</i></p> <p>Quiz 6 à faire (date limite 21 octobre, 23h59)</p>
	 Labo 6. À l'aide d'Excel et de SPSS, établir et visualiser la relation entre deux variables qualitatives au moyen de tableaux croisés et de diagrammes en barres groupées et empilées, les lire et interpréter en termes d'intensité et de direction de la relation. À l'aide de SPSS, visualiser des erreurs, les rechercher et les corriger par le biais de la fonction « rechercher et remplacer » de SPSS, manipuler la syntaxe de SPSS. Lire et interpréter les sorties SPSS, et présenter les résultats sous forme textuelle et dans des tableaux publiables à l'aide du logiciel Word.
	TP1 à rendre sur Studium (date limite 12 octobre)

	<i>Retour sur le TP1 et révision en vue de l'examen intra</i>
Semaine 7	
Vend. 21 oct.	Examen intra : Statistiques descriptives (séances 2, 3, 4, 5, 6) Local : B-3315 Pav. 3200 J. Brillant, Heure : 13h 00 à 16h 00.
Semaine 8	
Vend. 28 oct.	Période d'activités libres : Du 24 au 30 octobre
Semaine 9	
Vend. 04 nov.	 Leçon 7. Introduction à l'inférence statistique <ul style="list-style-type: none"> • Situations de recherche et conditions d'application de l'inférence statistique • Distinction entre échantillonnage aléatoire et échantillonnage empirique • Intelligence de l'inférence statistique : la distribution (théorie) d'échantillonnage • Estimation d'une moyenne et proportion d'une population (paramètres) à l'aide de la marge d'erreur et de l'intervalle de confiance • Logique des tests de validation d'hypothèses <p><i>Lectures suggérées : Fox, pp. 108-114 ; pp. 150-153</i></p> <p>Quiz 7 à faire (date limite 11 novembre, 23h59)</p>
	 Labo 7. À l'aide d'Excel et de SPSS, calculer l'erreur-type, la marge d'erreur et l'intervalle de confiance d'une statistique, les représenter graphiquement par la barre d'erreur afin de comparer des groupes eu égard à une variable quantitative. À l'aide de SPSS, éliminer des valeurs aberrantes par le biais de la sélection d'observations ou du recodage d'une variable, utiliser la syntaxe pour exécuter des instructions complexes avec SPSS. Lire et interpréter les sorties SPSS, présenter les résultats sous forme textuelle et dans des tableaux publiables à l'aide du logiciel Word. <p><i>Retour sur l'examen intra</i></p>
Semaine 10	
Vend. 11 nov.	 Leçon 8. Test du chi-carré <ul style="list-style-type: none"> • Situations de recherche et conditions d'application du test du chi-carré • Intelligence du chi-carré : simulation pour calculer les fréquences théoriques et comparaison avec les fréquences observées • Test du chi-carré sur des données de tableaux bivariés pour établir la signification statistique d'une relation entre deux variables qualitatives • Les mesures d'association pour établir la signification réelle et la direction d'une relation : Coefficient de contingence, V de Cramer et Gamma <p><i>Lectures suggérées : Fox, p.149 ; pp. 154-167 ; pp.169-175 ; pp.181-186</i></p> <p>Quiz 8 à faire (date limite 18 novembre, 23h59)</p>
	 Labo 8. À l'aide d'Excel, construire une table statistique de la distribution du chi-carré. À l'aide d'Excel et de SPSS, établir et visualiser la relation entre deux variables

	<p>qualitatives au moyen des tableaux croisés et diagrammes en bâtons groupés ou divisés, tester sa signification statistique (chi-carré) et mesurer sa signification réelle (V de Cramer) et direction (Gamma). Lire et interpréter les résultats en termes de signification statistique et réelle ainsi que du point de vue sociologique. Présenter les résultats sous forme textuelle et dans des tableaux publiables à l'aide du logiciel Word.</p> <p>TP2 à faire : Inférence statistique, test du chi-carré, test t, ANOVA</p>
Semaine 11	
Vend. 18 nov.	<p> Leçon 9. Comparaison de deux moyennes : test t</p> <ul style="list-style-type: none"> • Situations de recherche et conditions d'application du test t de <i>Student</i> • Tableau et diagramme des moyennes pour une relation entre une variable indépendante (VI) qualitative binaire et une variable dépendante (VD) quantitative • Intelligence du test de comparaison de deux moyennes : la différence observée et son erreur-type • Calcul et interprétation du test t en termes de signification statistique • Interprétation de l'éta-carré et du D de Cohen comme mesure d'association (signification réelle) <p><i>Lectures suggérées :</i> Fox, pp. 203-225</p> <p>Quiz 9 à faire (date limite 25 novembre, 23h59)</p>
	<p> Labo 9. À l'aide d'Excel, construire une table statistique du t de Student. À l'aide d'Excel et de SPSS, construire et interpréter des tableaux et diagrammes de moyennes ainsi que des diagrammes en boîtes juxtaposées illustrant la différence entre deux moyennes, tester cette différence au moyen du test t de Student, et en mesurer l'intensité au moyen de l'éta-carré et du D de Cohen. Lire et interpréter les résultats obtenus en termes de signification statistique et réelle ainsi que du point de vue sociologique. Présenter les résultats sous forme textuelle et dans des tableaux publiables à l'aide du logiciel Word.</p>
Semaine 12	
Vend. 25 nov.	<p> Leçon 10. Comparaison de trois moyennes ou plus : test ANOVA</p> <ul style="list-style-type: none"> • Situations de recherche et conditions d'application du test F d'ANOVA • Tableau et diagramme des moyennes pour une relation entre une VI qualitative non dichotomique et une VD quantitative • Intelligence du test de comparaison de trois moyennes ou plus : la variation intergroupes et la variation intragroupes • Calcul et interprétation du test F en termes de signification statistique • Interprétation de l'éta-carré comme mesure d'association (signification réelle) <p><i>Lectures suggérées :</i> Fox, pp. 227-252</p> <p>Quiz 10 à faire (date limite 2 décembre, 23h59)</p>

	 Labo 10. À l'aide d'Excel, construire une table statistique du F d'ANOVA. À l'aide d'Excel et de SPSS, construire et interpréter des tableaux et diagrammes de moyennes ainsi que des diagrammes en boîtes juxtaposées illustrant des différences entre trois moyennes ou plus, tester ces différences au moyen de l'ANOVA, en mesurer l'intensité au moyen de l'éta-carré. À l'aide de SPSS, déterminer les sources des différences au moyen des tests post hoc de comparaisons multiples (LSD, Bonferroni, Scheffe). Lire et interpréter les résultats obtenus en termes de signification statistique et réelle ainsi que du point de vue sociologique. Présenter les résultats sous forme textuelle et dans des tableaux publiables à l'aide du logiciel Word.
Semaine 13	
Vend. 02 déc.	 Leçon 11. Régression et corrélation linéaires <ul style="list-style-type: none"> • Situations de recherche et conditions d'application • Diagramme de dispersion pour un lien entre deux variables quantitatives • Intelligence de la modélisation linéaire : la droite et l'équation de régression • Calcul et interprétation de la constante et du coefficient de régression • Prédiction du score d'une VD connaissant le score d'une VI pour un cas • Calcul et interprétation du coefficient de corrélation et de son carré comme mesure d'association, test de signification statistique de la corrélation • Matrice de corrélations et gestion des données manquantes <p><i>Lectures suggérées : Fox, pp. 253-287</i></p> <p>Quiz 11 à faire (date limite 3 décembre, 23h59)</p>
	 Labo 11. À l'aide d'Excel et de SPSS, construire et interpréter des diagrammes de dispersion mettant en relation deux variables quantitatives, calculer le coefficient de corrélation linéaire et tester sa signification statistique à l'aide du test F, calculer l'équation de régression linéaire, procéder à la prédiction, construire une matrice de corrélations. Lire et interpréter les résultats en termes de prédiction, de signification statistique et réelle, ainsi que du point de vue sociologique. Présenter les résultats sous forme textuelle et dans des tableaux publiables à l'aide du logiciel Word. <p>TP2 à rendre (date limite 2 décembre, 23h59)</p>
Semaine 14	
Vend. 9 déc.	<p>À distance, sur Teams !</p> <p>Synthèse du cours et révision en vue de la préparation de l'examen final</p> <p>Retour sur le TP2</p>
Semaine 15	
Vend. 16 déc.	<p>Examen final : Statistiques inférentielles (séances 7, 8, 9, 10, 11)</p> <p>Local : B-3315 Pav. 3200 J. Brillant, Heure : 13h 00 à 16h 00.</p>

BIBLIOGRAPHIE

Volume de référence (obligatoire)

Fox, W. (1999). *Statistiques sociales*, Québec : Presses de l'université Laval (PUL), traduit de l'anglais et adapté par L.M. Imbeau (disponible à la librairie)

Références complémentaires (optionnels)

Imbeau, M.L. (2009). *Statistiques sociales avec SPSS. Cahier d'exercices*, 2^e éd Québec : PUL.

Lamoureux, A. (2003). *Recherche et méthodologie en sciences humaines*, Québec : Beauchemin.

Simard, C. (2008). *Méthodes quantitatives. Approche pédagogique progressive pour les sciences humaines*, 4^e édition.

Touré, El. H. (2011). *Le pouvoir des statistiques : informer ou contrôler?* Article. Notes de cours. Montréal : Université de Montréal. 2 p.

Touré, El. H. (2012). *Usages sociaux des sondages et inférence statistique. L'opinion publique introuvable.* Travail de réflexion critique. Notes de cours. Montréal : Université de Montréal. 5 p.