



[cursus]

LUC Adjengue

# Méthodes statistiques

Concepts, applications et exercices



## TABLE DES MATIÈRES

<b>AVANT-PROPOS</b>	<b>iii</b>
<b>LISTE DES FIGURES</b>	<b>xi</b>
<b>LISTE DES TABLEAUX</b>	<b>xv</b>
<b>1 PROBABILITÉS ET VARIABLES ALÉATOIRES</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Quelques définitions de base en probabilité . . . . .	1
1.2.1 Notions de base . . . . .	1
1.2.2 Probabilités . . . . .	3
1.2.3 Analyse combinatoire . . . . .	5
1.2.4 Probabilité conditionnelle et indépendance . . . . .	7
1.2.5 Axiome des probabilités totales et théorème de Bayes . . . . .	9
1.3 Les variables aléatoires . . . . .	10
1.3.1 Les variables aléatoires discrètes . . . . .	11
1.3.2 Les variables aléatoires continues . . . . .	11
1.3.3 La fonction de répartition . . . . .	12
1.3.4 Quelques caractéristiques d'une distribution . . . . .	13
1.3.5 Les fonctions d'une variable aléatoire . . . . .	15
1.3.6 La fonction génératrice des moments . . . . .	16
1.3.7 La fonction caractéristique d'une variable aléatoire . . . . .	16
1.3.8 Quelques lois de probabilité discrètes . . . . .	17
1.3.9 Quelques lois de probabilité continues . . . . .	22
1.4 Les vecteurs aléatoires . . . . .	35
1.4.1 Distribution conjointe d'un vecteur aléatoire . . . . .	35
1.4.2 Distributions marginales . . . . .	38
1.4.3 Indépendance stochastique . . . . .	39
1.4.4 Transformations de plusieurs variables . . . . .	40
1.4.5 Espérance mathématique et moments . . . . .	43
1.4.6 Fonction caractéristique d'un vecteur aléatoire . . . . .	47
1.4.7 Distributions conditionnelles . . . . .	48
1.4.8 La loi multinomiale . . . . .	51
1.4.9 Loi multinormale . . . . .	52

1.4.10 Le théorème central limite (cas multidimensionnel) . . . . .	55
1.5 Exercices . . . . .	56
<b>2 STATISTIQUE ET INFÉRENCE PARAMÉTRIQUE</b>	<b>67</b>
2.1 Introduction . . . . .	67
2.2 La statistique . . . . .	67
2.3 Statistique descriptive . . . . .	69
2.3.1 Représentations graphiques . . . . .	70
2.3.2 Description numérique . . . . .	72
2.3.3 Autres mesures et graphiques . . . . .	75
2.4 Échantillons aléatoires et lois d'échantillonnage . . . . .	82
2.4.1 Échantillons aléatoires . . . . .	82
2.4.2 Statistiques et distributions d'échantillonnage . . . . .	82
2.4.3 Lois d'échantillonnage usuelles . . . . .	85
2.5 Estimation . . . . .	94
2.5.1 Estimation ponctuelle : maximum de vraisemblance et moments . . . . .	94
2.5.2 Estimation par intervalle de confiance : cas d'un paramètre . . . . .	103
2.5.3 Estimation par intervalle de confiance : cas de deux paramètres . . . . .	109
2.6 Tests d'hypothèses paramétriques . . . . .	115
2.6.1 Hypothèses, tests, erreurs et risques . . . . .	116
2.6.2 Tests d'hypothèses pour une population . . . . .	118
2.6.3 Tests d'hypothèses pour deux populations . . . . .	129
2.6.4 Tests de normalité . . . . .	137
2.6.5 Le lemme de Neyman-Pearson et test du rapport de vraisemblances . . . . .	139
2.7 Exercices . . . . .	149
<b>3 STATISTIQUE ET TESTS NON PARAMÉTRIQUES</b>	<b>161</b>
3.1 Introduction . . . . .	161
3.2 Le test du khi-deux et ses applications . . . . .	161
3.2.1 La loi multinomiale . . . . .	162
3.2.2 Le principe général du test du khi-deux . . . . .	163
3.2.3 Le test d'ajustement : méthode du khi-deux . . . . .	165
3.2.4 Tableaux de contingence et test d'indépendance . . . . .	169
3.2.5 Le test d'homogénéité . . . . .	173
3.3 Autres tests non paramétriques . . . . .	178
3.3.1 Le test des signes . . . . .	178
3.3.2 Le test des rangs signés de Wilcoxon . . . . .	187
3.3.3 Le test de la somme des rangs de Wilcoxon (Mann -Whitney) . . . . .	191
3.3.4 Le test des rangs de Kruskal-Wallis . . . . .	195
3.3.5 Le test des rangs de Friedman . . . . .	199
3.3.6 Le coefficient de corrélation de Spearman . . . . .	202
3.4 Exercices . . . . .	204



<b>4</b>	<b>ANALYSE DE SÉRIES CHRONOLOGIQUES</b>	<b>215</b>
4.1	Introduction	215
4.2	Quelques définitions de base et modèles classiques	215
4.2.1	Les deux modèles	219
4.2.2	Identification du type de modèle	220
4.3	Décomposition d'une série chronologique	221
4.3.1	Analyse de la tendance d'une série chronologique	221
4.3.2	Analyse de la composante saisonnière	229
4.3.3	Désaisonnalisation d'une série chronologique	235
4.4	Quelques méthodes de calcul de prévision	236
4.4.1	Prévision par les moyennes mobiles	237
4.4.2	Prévision par décomposition	237
4.4.3	Prévision par lissage exponentiel simple	238
4.4.4	Prévision par lissage exponentiel avec tendance	240
4.4.5	Prévision par lissage avec tendance et composante saisonnière	241
4.5	Les modèles de Box et Jenkins	242
4.5.1	Convergence en moyenne quadratique	242
4.5.2	Processus stochastiques linéaires	243
4.5.3	Les modèles autorégressif, moyenne mobile et mixte	246
4.5.4	Estimation des autocorrélations et identification	256
4.5.5	Estimation des paramètres et validation du modèle	260
4.5.6	Calcul de prévisions	261
4.6	Exercices	266
<b>5</b>	<b>ANALYSE DE RÉGRESSION</b>	<b>271</b>
5.1	Introduction	271
5.1.1	Quelques exemples	272
5.1.2	Les principales étapes d'une analyse de régression	274
5.1.3	La classification des modèles de régression	275
5.2	Le modèle de régression linéaire simple	276
5.2.1	Estimation des paramètres	278
5.2.2	Inférence sur les paramètres $\beta_0$ et $\beta_1$	284
5.2.3	Analyse de la variance	287
5.2.4	Validation d'un modèle : diagnostics et mesures correctives	290
5.2.5	Utilisation de la droite de régression	302
5.3	La corrélation	305
5.3.1	Le coefficient de corrélation $\rho$	305
5.3.2	Le coefficient de corrélation $r$	305
5.3.3	L'inférence sur $\rho$	306
5.4	Le modèle de régression linéaire multiple	308
5.4.1	L'estimation des paramètres : méthode des moindres carrés	309
5.4.2	Propriétés des estimateurs	312
5.4.3	Estimation de la variance, intervalle de confiance et tests individuels	313
5.4.4	Tableau d'analyse de la variance et le test global	317
5.4.5	Utilisation de la courbe de régression	323
5.4.6	Vérifications sur la validité d'un modèle	325



5.5	La régression polynomiale . . . . .	327
5.6	Quelques aspects particuliers de la régression multiple . . . . .	329
5.6.1	Sommes de carrés et tests . . . . .	329
5.6.2	Sélection de variables . . . . .	340
5.6.3	Analyse de stabilité . . . . .	343
5.6.4	Utilisation de variables indicatrices . . . . .	348
5.7	La régression logistique . . . . .	355
5.7.1	Introduction et contexte . . . . .	355
5.7.2	Le modèle de régression logistique simple . . . . .	357
5.7.3	Le modèle de régression logistique multiple . . . . .	359
5.8	Quelques rappels sur les notions de vecteurs et de matrices . . . . .	360
5.8.1	Matrices et vecteurs . . . . .	360
5.8.2	Quelques opérations sur les matrices . . . . .	360
5.8.3	Quelques règles de calcul . . . . .	361
5.9	Exercices . . . . .	362
<b>6</b>	<b>ANALYSE DE VARIANCE</b>	<b>371</b>
6.1	Introduction . . . . .	371
6.2	Le modèle à un facteur . . . . .	372
6.2.1	Le modèle des moyennes . . . . .	373
6.2.2	Le modèle des effets . . . . .	373
6.2.3	La décomposition de la variation totale . . . . .	374
6.2.4	Le test de la neutralité des effets (égalité des moyennes) . . . . .	375
6.2.5	L'approche par l'analyse de régression . . . . .	380
6.2.6	L'estimation des paramètres, les intervalles de confiance et les tests . . . . .	381
6.2.7	L'analyse des résidus . . . . .	391
6.3	Le modèle à deux facteurs . . . . .	393
6.3.1	La planification d'expériences, les plans factoriels et les interactions . . . . .	393
6.3.2	L'analyse de variance du modèle à deux facteurs . . . . .	395
6.3.3	Les tests . . . . .	398
6.3.4	L'estimation des paramètres, les intervalles de confiance et les tests . . . . .	403
6.3.5	L'analyse des résidus . . . . .	410
6.4	Les plans d'expériences avec blocs . . . . .	411
6.4.1	Les blocs aléatoires complets . . . . .	411
6.4.2	La décomposition de la variation totale et les tests . . . . .	412
6.4.3	Les comparaisons multiples . . . . .	414
6.5	Les plans avec plusieurs facteurs : les plans $2^k$ . . . . .	415
6.5.1	Le cas général . . . . .	415
6.5.2	Le plan $2^2$ . . . . .	416
6.5.3	Les plans $2^k$ avec $k \geq 3$ . . . . .	420
6.5.4	Les blocs dans les plans $2^k$ . . . . .	423
6.6	Exercices . . . . .	426

<b>7 ANALYSE EN COMPOSANTES PRINCIPALES</b>	<b>433</b>
7.1 Introduction . . . . .	433
7.2 L'analyse en composantes principales . . . . .	434
7.2.1 Les données . . . . .	434
7.2.2 Les distances . . . . .	434
7.2.3 La visualisation des distances entre les points . . . . .	436
7.2.4 Les composantes principales . . . . .	440
7.3 L'utilisation de l'ACP en analyse de régression . . . . .	449
7.4 La détermination des composantes principales . . . . .	455
7.4.1 La première composante principale . . . . .	456
7.4.2 La deuxième composante principale . . . . .	456
7.4.3 La notation matricielle . . . . .	457
7.5 Exercices . . . . .	460
<b>A RÉPONSES AUX EXERCICES</b>	<b>463</b>
<b>B TABLES DES LOIS ET VALEURS CRITIQUES</b>	<b>489</b>
<b>C COURBES CARACTÉRISTIQUES</b>	<b>517</b>
<b>BIBLIOGRAPHIE</b>	<b>525</b>
<b>INDEX</b>	<b>527</b>



# Méthodes statistiques

Concepts, applications et exercices

*Méthodes statistiques – Concepts, applications et exercices* présente de façon synthétique un large éventail de méthodes statistiques rarement réunies dans un même ouvrage : inférence paramétrique, tests non paramétriques, analyse de séries chronologiques, analyse de régression, analyse de variance et analyse en composantes principales.

L'ouvrage s'ouvre sur un chapitre traitant de notions de probabilité, de variables et de vecteurs aléatoires ainsi que de lois de probabilité usuelles qui constitue une référence utile. De plus, il comporte de nombreux exemples et propose un grand nombre d'exercices résolus touchant des domaines variés. Lorsque nécessaire, particulièrement pour certaines analyses statistiques, l'auteur a recours à Matlab, à Excel, à Statistica et au logiciel libre R pour faciliter les calculs.

Ce livre s'adresse principalement aux étudiants universitaires de premier cycle en génie, en mathématiques, en gestion, en économie ou dans tout autre domaine exigeant une solide connaissance de base de la statistique.

**Luc Adjengue est professeur agrégé au Département de mathématiques et de génie industriel de l'École Polytechnique de Montréal. Il est détenteur d'un baccalauréat en statistique mathématique de l'Institut de statistique de Yaoundé, au Cameroun, ainsi que d'une maîtrise et d'un doctorat en statistique de l'Université de Montréal, au Canada. Il coordonne l'enseignement de cours de statistique de premier cycle universitaire et enseigne les séries chronologiques et les approches statistiques à la reconnaissance de formes aux cycles supérieurs.**

[pressespoly.ca](http://pressespoly.ca)



PRESSES  
INTERNATIONALES  
POLYTECHNIQUE

