



MÉTHODES STATISTIQUES EN GESTION

Michel Tenenhaus



DUNOD
ENTREPRISE

TABLE DES MATIÈRES

Avant-propos	1
Introduction	3
Chapitre 1 - Statistique descriptive	9
1. Les données	9
2. Description d'une variable qualitative	10
3. Description d'une variable quantitative	14
3.1 Mesures de la tendance centrale	14
3.2 Mesures de la dispersion	15
3.3 Visualisation de la dispersion	16
Chapitre 2 - Estimation et test	20
1. Échantillon représentatif	20
2. Estimation	21
2.1 Estimation d'une proportion p	22
2.2 Estimation d'une moyenne μ	23
2.3 Estimation d'une variance σ^2	24
3. Test d'hypothèse	25
3.1 La méthodologie du test d'hypothèse	26
3.2 Comparaison d'une proportion à un standard	27
3.3 Comparaison d'une moyenne à un standard	29
3.4 Test du Khi-deux d'indépendance	32
3.5 Test d'adéquation d'une loi de probabilité à des données	35
Chapitre 3 - Comparaison d'échantillons indépendants	38
1. Comparaison de deux échantillons indépendants	38
2. Comparaison de k échantillons indépendants	44
2.1 Formule de décomposition	45
2.2 Rapport de corrélation	46
2.3 Étude du test d'égalité des moyennes	46
2.4 Comparaisons multiples	48
2.5 Application	49

3.	Comparaison des variances	52
3.1	Comparaison de deux variances	53
3.2	Comparaison de deux moyennes lorsque les variances sont inégales	54
3.3	Comparaison de k variances	54

Chapitre 4 - La régression simple 56

1.	Aspects descriptifs de la régression simple	56
1.1	Le cas "Prix d'un appartement"	56
1.2	Droite des moindres carrés	59
1.3	Contribution de chaque observation à la construction de la droite des moindres carrés : le levier	62
1.4	Formule de décomposition	63
1.5	Coefficient de détermination R^2	64
1.6	Coefficient de corrélation R	65
2.	Aspects inférentiels de la régression simple	66
2.1	Le modèle statistique	66
2.2	Estimation des coefficients de régression a et b	66
2.3	Estimation de l'écart-type σ du terme résiduel	67
2.4	Intervalles de confiance de a et b	68
2.5	Tests d'hypothèse sur les coefficients de régression a et b	69
2.6	Analyse des résidus et mesure de l'influence des observations	72
2.7	Prévision	75
2.8	Élimination des observations atypiques	78

Chapitre 5 - La régression multiple 82

1.	Données et notations	82
2.	Le modèle statistique	83
3.	Les problèmes	83
4.	Estimation des paramètres du modèle	85
4.1	Estimation des coefficients de régression $\beta_0, \beta_1, \dots, \beta_k$	85
4.2	Estimation de l'écart-type σ du terme résiduel	86
5.	Intervalles de confiance des coefficients de régression	89
6.	Coefficient de détermination R^2 , R^2 -ajusté et corrélation multiple R	90
7.	Le test global	92
8.	Les tests marginaux	92
9.	Analyse des observations	94
9.1	Le levier	94
9.2	Analyse des résidus	95
9.3	Une mesure de l'influence d'une observation : le DFITS	95
10.	Intervalle de confiance de $E(Y_x)$ et de prévision de y_x	96
11.	Les méthodes de régression pas à pas	100
11.1	La régression pas à pas ascendante	100
11.2	La régression pas à pas descendante	102

12.	Étude du modèle simplifié	104
12.1	Comparaison entre le modèle complet et le modèle simplifié	104
12.2	Étude du modèle simplifié	105
13.	Analyse de la covariance	108

Chapitre 6 - L'analyse en composantes principales 115

1.	Analyse exploratoire de données multidimensionnelles	115
1.1	Étude descriptive des individus	116
1.2	Étude descriptive des variables	118
2.	L'analyse en composantes principales	121
2.1	Présentation de l'analyse en composantes principales suivant l'approche géométrique de Pearson	122
2.2	Présentation de l'analyse en composantes principales suivant l'approche de Hotelling	136
3.	La méthode de classification ascendante hiérarchique à l'aide du critère de Ward	138

Chapitre 7 - L'analyse factorielle des correspondances 145

1.	Le tableau de contingence	145
2.	Analyse exploratoire du tableau de contingence	147
2.1	Le tableau des fréquences	147
2.2	Le tableau des profils-lignes	148
2.3	Le tableau des profils-colonnes	150
2.4	Le tableau des attractions/répulsions	150
2.5	Le tableau des contributions au Khi-deux	152
3.	Présentation de l'analyse des correspondances	156
4.	Le nuage N(I) des profils-lignes	156
5.	L'analyse en composantes principales du nuage des profils-lignes	158
5.1	Étude de la première composante principale	159
5.2	Étude des autres dimensions	162
6.	L'analyse en composantes principales du nuage N(J) des profils-colonnes	164
7.	Lien entre les deux analyses	167
7.1	Les relations de transition	167
7.2	La formule de reconstitution des données	169
8.	Représentations graphiques	169
8.1	Représentation pseudo-barycentrique	170
8.2	Représentation barycentrique	173
9.	Étude de deux cas particuliers	176
9.1	L'effet Guttman	176
9.2	Tableau de contingence diagonal par blocs	179
10.	Étude de l'exemple "Élections présidentielles de 1981"	183

Chapitre 8 - L'analyse factorielle des correspondances multiples **197**

1. Données et notations	197
2. L'analyse des correspondances multiples	199
2.1 Propriétés du nuage $N(I)$ des profils-lignes	201
2.2 Propriétés du nuage $N(J)$ des profils-colonnes	201
2.3 Les relations de transition	202
2.4 Cosinus carrés des modalités	204
2.5 Contributions des modalités	204
2.6 Valeurs-tests	205
2.7 Formules de reconstitution	206
2.8 Représentations graphiques	207
2.9 L'analyse canonique généralisée	211
3. Analyse des correspondances multiples des caractéristiques physiques et psychiques des races canines	212
3.1 Analyse exploratoire des données	212
3.2 Résultats de l'analyse des correspondances multiples	214
4. Typologie des races canines	223
5. L'effet Guttman	229
5.1 Étude d'une échelle de Guttman parfaite à fréquences de profils égales	230
5.2 Traitement d'un exemple	232

Chapitre 9 - L'analyse discriminante **243**

1. Données et notations	243
2. Analyse descriptive des données	245
2.1 Pouvoir discriminant des variables	245
2.2 Les différents types de corrélation de l'analyse discriminante	247
3. L'analyse factorielle discriminante	249
3.1 Présentation de la méthode	249
3.2 Étude de l'exemple	252
4. L'analyse discriminante bayésienne	254
5. L'analyse discriminante pas à pas	256
5.1 Analyse discriminante pas à pas ascendante	257
5.2 Analyse discriminante pas à pas descendante	260
6. Étude du modèle simplifié	260
6.1 Analyse factorielle discriminante des variables Température, Soleil, Pluie	260
6.2 Analyse discriminante bayésienne des variables discriminantes Z'_1, Z'_2	261

Chapitre 10 - Décomposition d'une série chronologique **266**

1. Les données	266
2. Décomposition classique	267

3.	Le lissage polynomial	271
4.	Lissage robuste d'une série chronologique	273
4.1	Les concepts de base	273
4.2	Les lisseurs les plus courants	274
4.3	Lissage robuste de l'indice de la production industrielle de la France (hors BTP)	275
5.	Utilisation du programme Census-X11 pour désaisonnaliser une série chronologique	275
5.1	Notations et modèle	275
5.2	Les principes généraux	277
5.3	Procédure d'analyse d'une série	277

Chapitre 11 - La méthodologie de Box-Jenkins pour l'analyse et la prévision d'une série chronologique **285**

1.	La théorie	285
1.1	Les concepts de base	286
1.2	Processus autorégressif AR(p)	288
1.3	Processus moyenne mobile MA(q)	291
1.4	Processus autorégressif moyenne mobile ARMA(p, q)	295
1.5	Les processus non stationnaires non saisonniers	297
1.6	Les processus saisonniers	299
1.7	Prévision	300
1.8	Estimation	305
1.9	Validation	307
1.10	Modèle d'analyse d'interventions	309
2.	La pratique	311
2.1	Les ventes de champagne	311
2.2	Indice de la production industrielle de la France	319
	Annexe : calcul des espérances conditionnelles	329

Chapitre 12 - Les méthodes de lissage exponentiel **331**

1.	Le lissage exponentiel simple	332
2.	La méthode de Holt	326
3.	La méthode de Winters	340

Annexe 1. Rappels de calcul des probabilités **345**

1.	Les principales définitions	345
2.	Les tables statistiques	350

Annexe 2. Rappels d'algèbre linéaire **360**

Références **365**

Index **369**