
Table des matières

I	Programmation avec R	1
1	Introduction	3
1.1	Caractéristiques R	3
1.2	Installation de R	4
1.3	Démarrer et quitter une session R	4
1.4	Historique des fonctions et aide	6
1.5	Chargement et installation de packages	7
1.6	Environnements graphiques pour R	8
1.6.1	Le package <code>r-commander</code>	8
1.6.2	Le logiciel <code>RStudio</code>	8
1.7	Exécution en mode batch	8
1.8	Sites internet et référence officiels de R	10
2	Commandes de base	11
2.1	Les objets dans R	11
2.1.1	Vecteurs	12
2.1.2	Matrices et tableaux	22
2.1.3	Listes	24
2.1.4	Tableaux de données	25
2.1.5	Conversions	28
2.2	Valeurs spéciales	29
2.3	Opérations arithmétiques	30
2.4	Ordre de priorité	30
3	Gestion des fichiers sous R	33
3.1	Importation et exportation de données	33
3.1.1	Fichier au format texte	33
3.1.2	Fichier au format R	36
3.1.3	Fichier aux formats liés à d'autres logiciels	36

3.2	Exécution d'un script <code>R</code>	36
3.3	Redirection des sorties	36
3.4	Gestion des fichiers	37
4	Éléments de programmation	39
4.1	Opérateurs logiques	39
4.2	Instructions conditionnelles	40
4.3	Boucle <code>for</code>	41
4.4	Boucle <code>while</code>	42
4.5	Boucle <code>repeat</code>	42
4.6	Actions dans une boucle	43
4.7	Sélecteur <code>switch</code>	43
5	Fonctions	45
5.1	Construire des fonctions	45
5.2	Somme, minimum, etc.	48
5.3	Tri, rang et permutation	49
5.4	Comment appliquer une fonction à tout objet ?	49
5.5	Arrondis et troncatures	51
5.6	Mathématiques	52
5.6.1	Fonctions et opérateurs mathématiques	52
5.6.2	Méthodes numériques	53
6	Graphiques	59
6.1	Fonctions graphiques de haut niveau	59
6.1.1	Fonction <code>par</code>	59
6.1.2	Fonction <code>plot</code>	60
6.1.3	Fonction <code>curve</code>	65
6.1.4	Fonction <code>persp</code>	65
6.1.5	Fonction <code>matplot</code>	66
6.2	Fonctions graphiques de bas niveau	66
6.3	Partage de la fenêtre	70
6.3.1	Options graphiques <code>mfrow</code> et <code>mfc</code>	70
6.3.2	Fonction <code>split.screen</code>	71
6.3.3	Fonction <code>layout</code>	73
6.4	Les couleurs avec <code>R</code>	74
6.5	Exporter des graphiques	77
6.6	Le package <code>lattice</code>	79
6.7	Cartographies	80
6.7.1	Le package <code>maps</code>	80
6.7.2	Le package <code>maptools</code>	83

7	Création d'une interface graphique (GUI)	85
7.1	Le package <code>tcltk</code>	85
7.1.1	Simples fenêtres de dialogue	85
7.1.2	Des fenêtres plus complexes	86
7.2	Le package <code>tkrplot</code>	92
8	Déboguer un programme R	95
8.1	La fonction <code>traceback</code>	96
8.2	La fonction <code>debug</code>	97
8.3	La fonction <code>browser</code>	99
8.4	La fonction <code>trace</code>	100
9	Utiliser R avec d'autres langages	103
9.1	Intégrer des fonctions C dans un programme R	103
9.1.1	Étape 1 : écrire la fonction en C	103
9.1.2	Étape 2 : écrire le script R	104
9.1.3	Comparaisons	105
9.2	Utiliser des fonctions Fortran dans un programme R	106
9.2.1	Étape 1 : écrire la fonction en Fortran	106
9.2.2	Étape 2 : écrire le script R	107
9.2.3	Comparaisons	108
10	Création de package	111
10.1	Structure d'un package	111
10.2	La fonction <code>package.skeleton</code>	111
10.3	Vérification et création d'un package	112
II	Analyses statistiques avec R	115
11	Statistique descriptive	117
11.1	Statistiques univariées	117
11.1.1	Variables qualitatives	117
11.1.2	Variables quantitatives	121
11.2	Statistiques bivariées	129
11.2.1	Cas de deux variables qualitatives	129
11.2.2	Cas de deux variables quantitatives	132
11.2.3	Cas d'une variables quantitative et d'une variable qualitative	135
12	Lois de probabilité	139
12.1	Un peu de combinatoire	139
12.2	Quelques lois de probabilité	140
12.3	Simulations de variables aléatoires	143
12.4	Inférence statistique	144

12.4.1	Estimateur du maximum de vraisemblance pour quelques modèles paramétriques	144
12.4.2	Estimation non-paramétrique de la fonction de répartition . . .	150
12.4.3	Estimation non-paramétrique des quantiles	153
12.4.4	Estimation non-paramétrique de la densité	155
13	Quelques tests statistiques	165
13.1	Tests paramétriques de comparaison	165
13.1.1	Test d'égalité de proportions	166
13.1.2	Test d'égalité de moyennes de deux échantillons gaussiens . .	173
13.1.3	Test d'égalité de variances pour des échantillons gaussiens . .	178
13.1.4	Test de rapports de paramètres pour des lois de Poisson . . .	180
13.2	Tests non-paramétriques de comparaison	181
13.2.1	Test de comparaison de Kolmogorov-Smirnov	181
13.2.2	Test de Wilcoxon-Mann-Whitney	183
13.2.3	Test de Kruskal-Wallis	187
13.2.4	Diagramme quantiles-quantiles	188
13.3	Tests d'adéquation	189
13.3.1	Test pour la loi de Bernoulli	189
13.3.2	Test pour la loi de Poisson	189
13.3.3	Test d'adéquation du χ^2	190
13.3.4	Test d'adéquation de Kolmogorov-Smirnov	191
13.3.5	Test d'adéquation d'Anderson-Darling	193
13.3.6	Tests de normalité	193
13.3.7	Diagramme quantiles-quantiles	194
13.4	Tests d'indépendance et de corrélation	195
13.4.1	Test du χ^2 et variantes	195
13.4.2	Tests de corrélation	198
13.5	Tests de détection de valeurs aberrantes	201
13.5.1	Tests de Dixon	201
13.5.2	Tests de Grubbs	203
14	Modèles linéaires de régression	207
14.1	Régression linéaire	208
14.1.1	Régression linéaire simple	208
14.1.2	Régression linéaire multiple	224
14.1.3	Analyse des résidus	229
14.1.4	Outils de diagnostics	231
14.1.5	Comparaison et sélection de modèles	245
14.1.6	Transformation de Box-Cox	248
14.2	Analyse de la variance	251
14.2.1	Cas d'un seul facteur	252
14.2.2	Cas de deux facteurs	256
14.3	Analyse de la covariance	264
14.4	Régression logistique	267

15 Séries temporelles	271
15.1 Les classes spécifiques	271
15.1.1 La classe <code>Date</code>	271
15.1.2 La classe <code>ts</code>	272
15.2 Moyennes mobiles et lissages exponentiels	273
15.2.1 Moyennes mobiles (filtres linéaires)	273
15.2.2 Méthodes de lissage	275
15.2.3 Décomposition d'une série temporelle	278
15.3 Modèles SARIMA	279
15.3.1 Analyse préliminaire	281
15.3.2 Modèles autorégressifs	285
15.3.3 Modèles ARIMA	287
16 Analyse de durées de vie	293
16.1 La classe <code>Survival</code>	293
16.1.1 Durées censurées à droite	294
16.1.2 Durées censurées par intervalles	294
16.1.3 Durées censurées à droite et tronquées à gauche	295
16.2 Modèle paramétrique de vie accélérée	296
16.3 Méthodes non-paramétriques	298
16.4 Modèle à risques proportionnels et modèle de Cox	307
17 Analyse de données multivariées	309
17.1 Analyse en composantes principales	309
17.2 Analyse factorielle des correspondances	316
17.3 Analyse des correspondances multiples	322
18 Méthodes de classification	327
18.1 Classification ascendante hiérarchique	327
18.2 Méthode des k centres mobiles	332
19 Rapports automatiques	337
19.1 R et OpenOffice	337
19.2 R et \LaTeX	338
Index	343