

Table des matières

Avant-propos	1
Chapitre 1. Familles exponentielles	3
1.1. Définition	3
1.2. Moyenne, variance, fonction variance	4
1.3. Exemples de familles exponentielles	6
1.4. Maximum de vraisemblance	11
1.5. Annexe technique	20
1.5.1. Quelques résultats utiles de probabilités	20
1.5.2. Loi binomiale négative et mélange Poisson-gamma	21
1.6. Exercices	21
Chapitre 2. Du modèle linéaire aux GLM	27
2.1. Rappels sur le modèle linéaire	29
2.1.1. Écriture matricielle du modèle linéaire	30
2.1.2. Quelques exemples de modèles linéaires	31
2.1.3. Estimation par moindres carrés	33
2.1.4. Asymptotique de l'estimateur des moindres carrés	36
2.1.5. Modèle linéaire gaussien	39
2.1.5.1. Estimateur du maximum de vraisemblance	40
2.1.5.2. Lois des EMV	41
2.1.5.3. Intervalle de confiance pour un coefficient β_j	43
2.1.5.4. Test de Student de la nullité d'un coefficient β_j	43
2.1.5.5. Test de Student sur une combinaison linéaire des β_j	43
2.1.5.6. Test entre modèles emboîtés (test de Fisher)	44

2.2. Les trois composantes d'un modèle linéaire généralisé	45
2.2.1. Composante aléatoire	46
2.2.2. Prédicteur linéaire	47
2.2.3. Fonction de lien	47
2.3. Estimation dans les modèles linéaires généralisés	48
2.3.1. Maximum de vraisemblance	48
2.3.2. Propriétés asymptotiques et inférence	51
2.3.2.1. Intervalle de confiance, test de Wald, test du rapport de vraisemblance	52
2.3.3. Estimation du paramètre de dispersion	53
2.4. Quelques exemples	55
2.4.1. Le modèle de régression logistique	55
2.4.1.1. Cote et rapport de cote	59
2.4.1.2. Maximum de vraisemblance	60
2.4.2. Le modèle de régression de Poisson	62
2.4.3. Le modèle de régression gamma	63
2.5. Modèles linéaires généralisés avec R : exemple de la régression de Poisson	65
2.5.1. Intervalles de confiance et tests d'hypothèses	69
2.5.2. Critères AIC et BIC, sélection de variables	71
2.5.3. Prédiction, intervalle de confiance pour une prédiction	73
2.6. Annexe technique	75
2.6.1. Quelques lois de probabilité	75
2.6.2. Théorème de Cochran	75
2.7. Exercices	76

Chapitre 3. Données manquantes et censure dans les GLM 87

3.1. Données censurées	87
3.1.1. Introduction	87
3.1.2. Régression de Poisson avec réponse censurée à droite	90
3.1.2.1. Estimation	91
3.1.2.2. Quelques notations	93
3.1.2.3. Propriétés asymptotiques de l'EMV	94
3.1.2.4. Une petite étude de simulation	95
3.1.3. Régression gamma avec réponse censurée à droite	98
3.1.3.1. Quelques notations	100
3.1.3.2. Propriétés asymptotiques de l'EMV	101
3.1.3.3. Une petite étude de simulation	102
3.2. Problèmes à données manquantes	105
3.2.1. Introduction	105
3.2.2. Typologie des données manquantes	107
3.2.2.1. Données manquantes complètement au hasard	107

3.2.2.2. Données manquantes au hasard	107
3.2.2.3. Données manquantes non au hasard	108
3.2.3. Méthodes de traitement des données manquantes	108
3.2.3.1. Les différents cas	109
3.2.3.2. Pondération des cas complets	111
3.2.3.3. Méthodes d'imputation	113
3.2.3.4. Estimateur AIPW	115
3.2.4. Un problème de données manquantes dans le modèle de Poisson	118
3.2.4.1. Imputation par prédiction	119
3.2.4.2. Imputation multiple	125
3.2.4.3. Estimateur AIPW	128
3.2.5. Un problème de données manquantes en régression gamma	133
3.2.5.1. Estimateurs par imputation	134
3.2.5.2. Estimateur AIPW	138
3.3. Annexe technique	141
3.3.1. Deux lemmes	141
3.3.2. Preuve du théorème 3.3	146
3.3.2.1. Consistance	146
3.3.2.2. Normalité asymptotique	149
3.3.3. Preuve du théorème 3.4	150
3.3.4. Preuve du théorème 3.5	152
3.3.5. Preuve du théorème 3.6	156
3.3.6. Éléments de processus empiriques	157
3.3.6.1. Mesure, fonction de répartition et processus empiriques	157
3.3.6.2. Théorèmes de Glivenko-Cantelli et Donsker	158
3.3.6.3. Processus empirique indexé par des fonctions	159
3.3.6.4. Conditions suffisantes pour que \mathcal{F} soit Glivenko-Cantelli ou Donsker	160
3.4. Exercices	161

Chapitre 4. Modèles inflatés en zéro 167

4.1. Introduction	167
4.1.1. Surdispersion	167
4.1.2. Excès de zéros	171
4.2. Modèles de Poisson inflatés en zéro et extensions	174
4.2.1. Le modèle de Poisson inflaté en zéro	174
4.2.2. Les modèles ZIP semi-paramétriques	178
4.2.3. Le modèle de Poisson généralisé inflaté en zéro	182
4.2.3.1. La loi de Poisson généralisée	182
4.2.3.2. Le modèle ZIGP	183
4.2.3.3. Propriétés asymptotiques de l'EMV	184
4.2.4. Un test d'excès de zéros	185

- 4.3. Le modèle binomial négatif inflaté en zéro 190
 - 4.3.1. Le modèle binomial négatif 190
 - 4.3.2. Le modèle ZINB 192
 - 4.3.3. Modèle ZIP *versus* modèle ZINB 194
- 4.4. Le modèle binomial inflaté en zéro 195
- 4.5. Censure et données manquantes : exemples de problèmes 197
 - 4.5.1. Le modèle ZIP censuré 198
 - 4.5.2. Covariables manquantes dans le modèle ZIB 200
 - 4.5.3. Covariables manquantes dans le modèle ZIP 203
 - 4.5.3.1. Cas où les covariables observées sont catégorielles 203
 - 4.5.3.2. Cas où des covariables observées sont continues 206
- 4.6. Modèles inflatés en zéro marginaux 208
 - 4.6.1. Introduction 208
 - 4.6.2. Modèles MZIP et MZINB 210
- 4.7. Exercices 212

Notations et sigles 217

Bibliographie 219

Index 229