

collection

L1 SANTÉ

Statistiques médicales et biologiques

Fethi BORSALI



SOMMAIRE

Partie 1 : STATISTIQUE DESCRIPTIVE

- Chapitre 1 : Etude univariée	12
I. Généralités et définitions	12
1. Le recueil des données	12
2. La présentation des données	13
3. L'analyse des données	13
4. La fiabilité des résultats	13
II. Présentation numérique et graphique	13
1. Caractère quantitatif discret	13
2. Caractère quantitatif continu	16
3. Caractère qualitatif	19
III. Paramètres caractéristiques	22
1. Paramètres de position	22
1.1. Moyenne arithmétique	22
1.2. Mode	24
1.3. Médiane	24
1.4. Quartiles	25
1.5. Déciles	27
1.6. Centiles	27
1.7. Représentation graphique « Box Plot »	27
2. Paramètres de dispersion	30
2.1. Variance	30
2.2. Ecart-type	30
2.3. Signification de la variance (ou de l'écart-type)	31
2.4. Ecart-interquartile	32
2.5. Coefficient de dispersion	32
3. Paramètres de forme	32
3.1. Moments	32
3.2. Coefficient d'asymétrie	33
3.3. Coefficient de dérive	34
3.4. Coefficient d'aplatissement	34
IV. Conclusion	35
- Chapitre 2 : Etude bivariée	36
I. Introduction	36
II. Ajustement linéaire	37
1. Droite de régression de y par rapport à x	38
2. Droite de régression de x par rapport à y	38
3. Covariance	39
4. Détermination des coefficients	39
5. Coefficient de corrélation de PEARSON	40
6. Interprétation concrète du coefficient de corrélation	40

7. Coefficient de corrélation et type de corrélation.....	42
III. Ajustement non linéaire.....	45
IV. Représentation par un tableau de contingence.....	47
V. Conclusion	51
 Partie 2 : STATISTIQUE PROBABILISTE	
CHAPITRE 1 : Eléments d'analyse combinatoire et calcul de probabilités.....	52
 I. Analyse combinatoire	
1. Généralités.....	52
2. Arrangements	53
2.1. Arrangements sans répétition	53
2.2. Arrangements avec répétition	53
3. Permutations.....	54
3.1. Permutations sans répétition.....	54
3.2. Permutations avec répétition	54
4. Combinaisons.....	54
4.1. Combinaisons sans répétition.....	54
4.2. Combinaisons avec répétition	55
5. Binôme de NEWTON	56
 II. Calcul de probabilités	
1. Généralités et définitions.....	57
2. Notion d'événement.....	57
3. Définition d'une probabilité	58
3.1. Définition classique	58
3.2. Définition statistique	59
3.3. Remarque importante	59
3.4. Quelques exemples	59
4. Opérations sur les événements.....	60
5. Axiomes des probabilités.....	61
6. Probabilités composées.....	62
6.1. Evénements dépendants. Probabilités conditionnelles.....	62
6.2. Evénements indépendants	63
7. Probabilités totales générales	64
8. Probabilités complètes	64
9. Probabilités des causes	66
 Chapitre 2 : Variables aléatoires réelles.....	
 I. Introduction et définition	
1. Exemple préliminaire	69
2. Définition	70
 II. Variables aléatoires discrètes	
1. Définition	70
2. Loi de probabilité.....	70
3. Fonction de distribution	71
 III. Variables aléatoires continues	
1. Définition	72

2. Loi de probabilité.....	72
3. Fonction de distribution	72
4. Fonction densité ou densité de probabilité	72
IV. Caractéristiques d'une variable aléatoire	74
1. Espérance mathématique.....	74
2. Variance	75
3. Ecart-type	75
4. Covariance	76
V. Inégalité de BIENAYME-TCHEBICHEFF	76
1. Interprétation de l'inégalité de BIENAYME-TCHEBICHEFF	77
2. Signification de la variance (ou de l'écart-type)	77
Chapitre 3 : Lois de probabilité fondamentales	79
I. Généralités et définitions	79
II. Loi binomiale	79
1. Définition	79
2. Espérance mathématique et variance.....	80
3. Généralisation de la loi binomiale : loi multinomiale	81
4. Loi faible des grands nombres de JACOB BERNOULLI	81
III. Loi hypergéométrique.....	83
1. Définition	83
2. Espérance mathématique et variance.....	83
3. Approximation d'une loi hypergéométrique à une loi de binomiale ..	83
IV. Loi de POISSON	83
1. Phénomènes qui suivent la loi de POISSON.....	84
2. Définition	85
3. Espérance mathématique et variance.....	85
4. Approximation d'une loi binomiale à une loi de POISSON	85
V. Loi exponentielle	86
1. Définition	86
2. Espérance mathématique et variance.....	86
3. Phénomènes qui suivent la loi exponentielle	86
4. fonction de survie.....	87
VI. Loi de LAPLACE-GAUSS ou loi normale	87
1. Phénomènes qui suivent la loi normale	87
2. Définition	92
3. Espérance mathématique et variance.....	92
4. Etude des caractères biologiques	92
5. Loi normale centrée réduite	93
6. Utilisation de la table	94
7. Propriétés de la loi normale.....	96
8. Approximation d'une loi binomiale à une loi normale.....	97
Partie 3 : STATISTIQUE INDUCTIVE	
Chapitre 1 : Jugement sur échantillons	99
I. Introduction.....	99
II. Fluctuations d'échantillonnage	100

III.	Interprétation statistique	100
1.	Les problèmes d'estimation	100
2.	Les problèmes de conformité	101
3.	Les problèmes d'homogénéité	101
4.	Les problèmes de relation	102
IV.	Remarque sur la distribution du caractère	103
V.	Conclusion	103
Chapitre 2 : Echantillonnage et estimation		104
I.	Généralités et définitions	104
II.	Distribution d'échantillonnage et intervalle de confiance d'une moyenne	104
A.	Cas des grands échantillons ($n \geq 30$)	104
1.	Distribution d'échantillonnage d'une moyenne	104
2.	Intervalle de confiance d'une moyenne	106
3.	Précision de l'estimation	107
4.	Détermination du nombre d'individus nécessaire	108
B.	Cas des petits échantillons ($n < 30$)	109
1.	Distribution d'échantillonnage d'une moyenne	109
2.	Intervalle de confiance d'une moyenne	110
III.	Distribution d'échantillonnage et intervalle de confiance d'une proportion	112
1.	Distribution d'échantillonnage d'une proportion	112
2.	Intervalle de confiance d'une proportion	114
3.	Précision de l'estimation	115
4.	Détermination du nombre d'individus nécessaire	116
Chapitre 3 : Théorie des tests statistiques		118
I.	Position du problème	118
II.	Erreurs et risques d'erreurs	119
III.	Degré de signification	123
IV.	Puissance du test	124
1.	Influence de l'hypothèse alternative	124
2.	Influence de la taille de l'échantillon	125
V.	Détermination du nombre de sujets nécessaire	126
1.	Comparaison entre deux moyennes	126
2.	Comparaison entre deux proportions	127
VI.	Inflation du risque d'erreur de première espèce	128
Chapitre 4 : Comparaison de moyennes et de proportions		131
I.	Comparaison de moyennes	131
1.	Comparaison d'une moyenne observée à une moyenne théorique ..	131
2.	Comparaison de deux moyennes pour des observations indépendantes	133
3.	Comparaison de deux moyennes pour des observations appariées ..	141
II.	Comparaison de deux proportions	145

1. Comparaison d'une proportion observée à une proportion théorique	145
2. Comparaison de deux proportions pour des observations indépendantes.....	147
3. Comparaison de deux proportions pour des observations appariées.....	150
Chapitre 5 : Analyse de la variance (ANOVA).....	157
I. Introduction	157
II. Analyse de la variance à un facteur (ANOVA à un facteur)	157
1. Définitions.....	158
2. Décomposition de la variance totale.....	159
3. Estimation de la variance totale.....	159
4. Test <i>F</i> de FISHER-SNEDECOR	160
5. Analyse des comparaisons a posteriori -Méthode de SCHEFFE-	162
6. Taille de l'effet et la statistique éta au carré (η^2)	170
III. Analyse de la variance à deux facteurs (ANOVA à deux facteurs)	170
1. Plan d'ordre deux à une répétition	170
1.1. Position du problème	171
1.2. Décomposition de la variance totale	173
1.3. Test <i>F</i> de FISHER-SNEDECOR.....	173
2. Plan d'ordre deux à plusieurs répétitions	178
2.1. Position du problème	178
2.2. Décomposition de la variance totale	181
2.3. Test <i>F</i> de FISHER-SNEDECOR.....	181
IV. Analyse des comparaisons a priori dans l'ANOVA à un facteur	184
1. Contrastes orthogonaux.....	185
2. Inflation du risque d'erreur.....	186
3. Analyse des comparaisons a priori	186
3.1. Contrastes orthogonaux	188
3.2. Contrastes orthogonaux	188
V. Conclusion	191
Chapitre 6 : Test de khi-deux ou test de K. PEARSON.....	192
I. Comparaison d'une distribution expérimentale a une distribution théorique	192
II. Comparaison simultanée de plusieurs distributions expérimentales.....	196
III. Comparaison de deux caractères qualitatifs	202
Chapitre 7 : Méthodes d'ajustement	209
I. Introduction	209
II. Comparaison de deux proportions	209
1. Méthode de MANTEL-HAENSZEL.....	209
2. Méthode de COCHRAN.....	212
III. Comparaison de deux moyennes	214
Chapitre 8 : Introduction à l'épidémiologie	215

I.	Introduction	215
II.	Risque et facteurs de risque	215
III.	Association statistique et causalité	215
1.	Association fallacieuse	216
2.	Association inverse	216
3.	Association indirecte	216
IV.	Organisation et résultats des enquêtes	217
1.	Enquêtes sur échantillon représentatif « Cross-sectional study »	218
2.	Enquêtes exposé-non exposé ou enquête prospective « Cohort study » ou « Follow-up study »	219
3.	Enquêtes cas-témoin ou enquête rétrospective « Case-control study »	220
4.	Enquêtes croisée « Cross-over study »	221
5.	Essai contrôlé randomisé « Randomized controlled trial »	222
V.	Risque attribuable	225
1.	Fraction étiologique chez les exposés	225
2.	Fraction étiologique dans la population	225
VI.	Analyse de stratification	225

Chapitre 9 : Plan factoriel 2x2 230

I.	Introduction	230
II.	Interaction entre deux traitements	230
1.	Organisation de l'essai	230
2.	Analyse des données	230
181	2.1. Résultat qualitatif à deux classes : succès ou échec	230
182	2.2. Résultat quantitatif	234
III.	Interaction ordre et traitement	236
183	1. Organisation de l'essai	236
184	2. Analyse des données	236

Chapitre 10 : Evaluation des méthodes diagnostiques 244

I.	Introduction	244
II.	Classification des sujets T_+ et T_- entre les groupes des malades et non malades	245
III.	Estimation des probabilités	246
1.	Echantillon représentatif	246
2.	Echantillon de sujets présentant ou non le signe	247
3.	Echantillon de malades et de non malades	247
IV.	Erreurs et risques d'erreurs	248
V.	Rapports de vraisemblances	248
VI.	Evaluation des examens diagnostiques dont le résultat est quantitatif	251
1.	Construction de la courbe R.O.C	251
2.	Interprétation de la courbe R.O.C	251

Chapitre 11 : Analyse de survie 256

I.	Introduction	256
II.	Méthodes d'analyses	256

1. Méthode de KAPLAN-MEIER	257
2. Courbe de survie	259
3. Méthode actuarielle	260
4. Courbe de survie	262
III. Comparaison de plusieurs courbes de survie	265
1. Méthode de LOGRANK	265
2. Méthode de LOGRANK ajustée	271
Chapitre 12 : Corrélation et régression linéaire simple.....	281
I. Corrélation simple	281
II. Corrélation partielle	284
III. Régression linéaire simple.....	285
1. Corrélation et régression	285
2. Qualité de l'ajustement : pourcentage de variance expliquée par la régression	286
3. Test de la pente de la droite de régression par l'analyse de la variance.....	287
4. Analyse des résidus	288
IV. Test de la linéarité de la régression	288
V. Epreuve graphique de la normalité d'une distribution et détermination rapide de la moyenne et de la variance	295
Chapitre 13 : Régression linéaire multiple.....	298
I. Introduction	298
II. Evaluation globale du modèle	299
III. Evaluation individuelle des coefficients de régression.....	301
IV. Coefficient de détermination R^2 et coefficient de corrélation R	302
V. Analyse des résidus : validation des hypothèses du modèle	302
VI. Traiterhent des données aberrantes (atypiques).....	308
VII. Sélection des variables explicatives.....	309
1. Procédure ascendante (forward selection).....	309
2. Procédure descendante (backward elimination)	309
3. Procédure pas à pas (stepwise selection)	309
VIII. Evaluation d'un modèle réduit	310
IX. Evaluation d'un modèle avec quelques variables explicatives fixées	310
X. Conclusion	311
Chapitre 14 : Analyse de covariance (ANCOVA)	312
I. Introduction	312
II. Analyse de covariance	313
1. Régressions linéaires simples	313
2. Parallélisme des droites de régression	314
3. Ordonnées à l'origine	315
III. interprétation graphique de l'analyse de covariance.....	322
IV. Conditions d'utilisation de l'analyse de covariance	325
Chapitre 15 : Régression logistique.....	327

I.	Introduction	327
II.	Régression logistique simple	327
III.	Régression logistique multiple	328
IV.	Etude des données	328
V.	Estimation des paramètres du modèle	329
VI.	Evaluation des coefficients de régression	330
1.	Régression logistique simple	330
2.	Régression logistique multiple	330
VII.	Qualité de l'ajustement du modèle	331
VIII.	Sélection des variables explicatives	333
1.	Procédure ascendante (forward selection)	333
2.	Procédure descendante (backward elimination)	334
3.	Procédure pas à pas (stepwise selection)	335
IX.	Analyse des résidus	336
X.	Interprétation des coefficients	336
1.	Rappels sur le odds ratio <i>OR</i>	337
2.	Coefficient de régression d'une variable explicative dichotomique	337
3.	Coefficient de régression d'une variable explicative quantitative	338
4.	Coefficient de régression d'une variable explicative polytomique	338
5.	Coefficient de régression de plusieurs variables explicatives	340
5.1.	Cas d'un modèle non additif	340
5.2.	Cas d'un modèle additif	341
XI.	Conclusion	342

Chapitre 16 : Analyse en composantes principales (A.C.P) 343

I.	Introduction	343
II.	Analyse des données	348
1.	Réduction des données	348
2.	Projection des points	349
3.	Qualité de l'ACP	351
4.	Contribution et qualité de la représentation	351
5.	Intérêt de l'ACP	356
5.1.	Test de sphéricité de BARTLETT	356
5.2.	Test KAISER-MAYER-OLKIN	357
6.	Conclusion	358

Chapitre 17 : Analyse factorielle des correspondances (A.F.C) 359

I.	Introduction	359
II.	Analyse des données	361
1.	Tableaux de contingences	361
2.	Distance du khi-deux	362
3.	Principaux indicateurs statistiques utilisés	364
3.1.	Taux de liaison entre deux caractères	364
3.2.	Variance de la distribution des observations ou phi-deux de K. PEARSON (Φ^2)	364
4.	Méthode d'analyse	364
4.1.	Propriétés	364

4.2. Présentation géométrique des données (sur 2 axes factoriels)	365
4.3. Principaux éléments constitutifs de l'AFC	366
5. Interprétation des données	367
Chapitre 18 : Analyse factorielle des correspondances multiples (A.F.C.M)	368
I. Introduction	368
II. Présentation des données	368
1. Tableau initial.....	368
2. Tableau d'effectifs	369
3. Tableau disjonctif complet	369
4. Tableau disjonctif des patrons	369
5. Tableau de BURT	370
III. Analyse factorielle des correspondances multiples	371
IV. Analyse et interprétation.....	371
1. Distances entre profils lignes.....	371
2. Distances entre profils colonnes	371
3. Phi-deux de K. PEARSON	372
4. Conclusion	372
Chapitre 19 : Tests non paramétriques	373
I. Intérêt des tests non paramétriques	373
II. Traitement des ex-æquo	373
III. Tests non paramétriques	373
1. Statistique U de MANN-WHITNEY	374
2. Statistique W de WILCOXON	375
3. Statistique H de KRUSKALL-WALLIS	376
4. Statistique S de FRIEDMAN.....	379
5. Coefficient de corrélation de SPEARMAN ou rho de SPEARMAN : mesure de corrélation	382
6. Coefficient KAPPA : mesure de concordance entre observateurs.....	384
7. Statistique exact de FISHER	386
8. Statistique de KOLMOGOROV-SMIRNOV : ajustement des données à la loi normale	388
IV. Conclusion	390
TABLES NUMERIQUES	391

Le présent ouvrage est tout particulièrement destiné aux étudiants en médecine, en pharmacie, en chirurgie dentaire et en biologie, ainsi qu'aux chercheurs des disciplines médicales et biologiques désirant compléter leur travail de recherche par des méthodes statistiques ou s'initier à des méthodes statistiques plus approfondies.

L'auteur a voulu exposer, de manière élémentaire, le programme de statistique des premières années de médecine, de pharmacie, de chirurgie dentaire et de biologie (statistique descriptive, statistique probabiliste et les six premiers chapitres de la statistique inductive) et approfondir suffisamment ces programmes pour permettre aux chercheurs d'effectuer le traitement des données recueillies au cours d'observations de phénomènes (à partir du septième chapitre de la statistique inductive). Chaque chapitre de cet ouvrage contient de nombreux exemples basés sur des situations rencontrées dans la discipline médicale et biologique. Certaines notions ont été détaillées jusqu'à déborder légèrement le programme. La plupart des démonstrations ont été évitées car elles nécessitent des connaissances mathématiques très profondes que l'on ne peut exiger des étudiants en médecine, en pharmacie, en chirurgie dentaire et en biologie. La lecture de cet ouvrage ne nécessite aucune connaissance de mathématiques à l'exception de quelques notions très simples de l'enseignement secondaire.

Fethi Borsali enseigne les biostatistiques à la faculté d'Oran.



9 782729 853457

www.editions-ellipses.fr