

# Sommaire

AVERTISSEMENT .....	XI	4. Mode .....	30		
		5. Moyenne .....	30		
		6. Fréquence relative .....	33		
		7. Pourcentage .....	33		
<b>PARTIE I</b>					
<b>STATISTIQUE DESCRIPTIVE</b>					
<b>CHAPITRE 1 – VARIABLES</b>					
<b>I. Variables quantitatives</b> .....	3	<b>II. Paramètres de dispersion</b> .....	34		
1. Variables quantitatives continues .....	3	1. Extrêmes .....	34		
2. Variables quantitatives discrètes .....	4	2. Étendue .....	34		
3. Variables temporelles .....	4	3. Intervalle interquartile et semi-interquartile .....	34		
<b>II. Variables qualitatives</b> .....	5	4. Variance .....	35		
1. Variables qualitatives ordinales .....	5	5. Écart type .....	36		
2. Variables qualitatives nominales ou catégorielles .....	5	6. Coefficient de variation .....	36		
3. Variables binaires .....	5	7. Variance et écart type d'une variable qualitative binaire .....	37		
<b>CHAPITRE 2 – ORGANISATION DES DONNÉES</b>					
<b>I. Tri des données</b> .....	11	<b>CHAPITRE 5 – REPRÉSENTATION D'UNE DISTRIBUTION</b>			
<b>II. Regroupement en classes</b> .....	11	<b>I. Variable discrète :</b> <b>fréquences relatives des classes</b> .....	42		
<b>III. Transformation de variable</b> .....	13	<b>II. Variable continue :</b> <b>densité de probabilité</b> .....	42		
<b>IV. Effectifs et fréquences</b> .....	15	<b>III. Symétrie et étalement d'une distribution</b> .....	44		
<b>V. Distribution</b> .....	16	1. Coefficient de dissymétrie (ou d'asymétrie) $\gamma_1$ ou <i>skewness</i> .....	44		
<b>CHAPITRE 3 – DESCRIPTION DES DONNÉES</b>				2. Coefficient d'aplatissement $\gamma_2$ ou <i>kurtosis</i> .....	44
<b>I. Tableaux</b> .....	17	<b>IV. Cas d'une variable qualitative binaire</b> .....	45		
1. Tableau brut de données .....	17	<b>CHAPITRE 6 – LOIS DE DISTRIBUTION</b>			
2. Tableaux de fréquences .....	18	<b>I. Loi binomiale</b> .....	47		
3. Problème des données manquantes .....	20	1. À quoi sert la loi binomiale ? .....	47		
<b>II. Graphiques</b> .....	20	2. Définition des termes de la loi binomiale .....	48		
1. Polygone de fréquence .....	21	3. Conditions d'application de la loi binomiale .....	50		
2. Histogramme .....	22	4. Propriétés additives de la loi binomiale .....	51		
3. Diagramme en barres .....	23	5. À quoi sert la loi binomiale ? (bis) .....	51		
4. Diagramme en barres horizontales .....	23	<b>II. Loi de Poisson</b> .....	52		
5. Camembert .....	23	1. À quoi sert la loi de Poisson ? .....	53		
6. Pyramide .....	24	2. Conditions d'application de la loi de Poisson .....	54		
<b>CHAPITRE 4 – MESURES EN STATISTIQUE</b>				<b>III. Loi normale</b> .....	55
<b>I. Paramètres de position</b> .....	27	1. Propriétés de la loi normale .....	56		
1. Médiane .....	28	2. Loi normale cumulée .....	56		
2. Quartiles .....	29	3. Loi normale centrée réduite .....	57		
3. Déciles et percentiles .....	29	4. Propriétés de la loi de Z normale centrée réduite .....	58		

## PARTIE II ESTIMATION

### CHAPITRE 7 – SONDAGE

<b>I. Biais de sélection</b>	.....	65
<b>II. Tirage au sort: le hasard</b>	.....	66
<b>III. Sondages aléatoires</b>	.....	66
1. Sondage élémentaire	.....	67
2. Sondage systématique	.....	67
3. Sondage à plusieurs degrés	.....	69
4. Sondage en grappes	.....	70
5. Sondage stratifié	.....	70
6. Sondages stratifiés à plusieurs degrés	.....	71
<b>IV. Sondages empiriques</b>	.....	71
1. Méthode des quotas	.....	71
2. Méthode des itinéraires	.....	72
3. Méthode des transects	.....	72
4. Méthode des unités-types	.....	72

### CHAPITRE 8 – MESURES STATISTIQUES SUR UN ÉCHANTILLON

<b>I. Paramètres de position</b>	.....	75
1. Moyenne	.....	75
2. Pourcentage	.....	75
<b>II. Paramètres de dispersion</b>	.....	76
1. Variance	.....	76
2. Écart type	.....	76

### CHAPITRE 9 – ESTIMATION D'UN PARAMÈTRE

<b>I. Estimation d'une moyenne inconnue</b>	.....	78
1. Fluctuation d'échantillonnage d'une moyenne	.....	78
2. Écart type de la moyenne	.....	79
3. Intervalle de confiance d'une moyenne	.....	79
4. Signification de l'intervalle de confiance d'une moyenne	.....	80
<b>II. Estimation d'un pourcentage inconnu</b>	.....	80
1. Fluctuation d'échantillonnage d'un pourcentage	.....	80
2. Écart type d'un pourcentage	.....	81
3. Intervalle de confiance d'un pourcentage	.....	81
4. Signification de l'intervalle de confiance d'un pourcentage	.....	82
<b>III. Risque d'erreur consentie <math>\alpha</math></b>	.....	82
<b>IV. Taille d'un échantillon</b>	.....	83
1. Précision d'une estimation	.....	83
2. Calcul de la taille d'un échantillon	.....	83

## PARTIE III TESTS ET MODÈLES STATISTIQUES

### INTRODUCTION

Conditions d'utilisation d'un test	.....	89
Conditions d'application	.....	90

### CHAPITRE 10 – PRINCIPE DES TESTS

<b>I. Principe des tests de comparaison</b>	.....	91
1. Établir l'hypothèse nulle ( $H_0$ )	.....	91
2. Proposer une hypothèse alternative ( $H_1$ )	.....	92
3. Calcul d'un test de comparaison	.....	92
4. Résultats d'un test de comparaison	.....	93
5. Choix du risque d'erreur	.....	94
6. Interprétation finale d'un test de comparaison	.....	95
<b>II. Principe des tests de liaison</b>	.....	97

### CHAPITRE 11 – TESTS DE COMPARAISON

<b>I. Test Z ou test de l'écart réduit</b>	.....	101
1. Principe du test Z	.....	101
2. Interprétation du test Z avec un risque $\alpha$ fixé à 5 %	.....	102
3. Utilisation pratique de la table Z	.....	102
4. Calcul du nombre de sujets nécessaires à un test Z	.....	104
<b>II. Test T de Student</b>	.....	105
1. Principe du test T	.....	105
2. Interprétation du test T	.....	105
3. Utilisation de la table T	.....	106
<b>III. Test F de Fisher-Snedecor</b>	.....	106
1. Principe du test F de comparaison de deux variances	.....	107
2. Analyse de la variance pour comparer plusieurs moyennes	.....	107
3. Utilisation des tables de F	.....	108
<b>IV. Tests de <math>\chi^2</math></b>	.....	110
1. Principe du $\chi^2$	.....	110
2. Interprétation du test $\chi^2$ avec un risque $\alpha$ fixé à 5 %	.....	113
3. Utilisation de la table de $\chi^2$	.....	114
<b>V. Test exact de Fisher</b>	.....	114
1. Principe du test exact de Fisher pour comparer deux pourcentages	.....	115
2. Calcul du test de Fisher	.....	115
<b>VI. Tests non paramétriques ou tests de rangs</b>	.....	116

**CHAPITRE 12 – TESTS DE LIAISON**

<b>I. Test du <math>\chi^2</math> d'indépendance</b> .....	119
1. Principe .....	119
2. Interprétation du test du $\chi^2$ d'indépendance ....	120
<b>II. Test du <math>\chi^2</math> de tendance</b> .....	120
<b>III. Tests de corrélation</b> .....	121
1. Covariance.....	122
2. Coefficient de corrélation .....	122
3. Test du coefficient de corrélation.....	123
4. Test de corrélation des rangs de Spearman .....	123
<b>IV. Régression linéaire simple</b> .....	124
1. Description .....	124
2. Test de la pente de la droite de régression.....	126
3. Estimations.....	126

**CHAPITRE 13 – UTILISATION PRATIQUE DES TESTS STATISTIQUES**

<b>I. Critères de choix d'un test statistique</b> .....	129
<b>II. Stratégie d'utilisation des tests statistiques</b> .....	130
1. Domaines d'application .....	130
2. Choix d'un test en fonction de la nature du problème .....	130
3. Choix d'un test en fonction des paramètres à comparer .....	132
4. Conditions d'application des tests .....	133
<b>III. Test Z pour comparer une moyenne observée à une moyenne théorique</b> .....	134
<b>IV. Test Z pour comparer deux moyennes</b> .....	136
<b>V. Test Z pour comparer deux moyennes sur deux séries appariées</b> .....	138
<b>VI. Test T pour comparer une moyenne observée à une moyenne théorique</b> .....	140
<b>VII. Test T de Student pour comparer deux moyennes</b> .....	142
<b>VIII. Test T pour comparer 2 moyennes sur 2 séries appariées</b> .....	144
<b>IX. Test F pour comparer deux variances</b> .....	146
<b>X. Test F pour comparer plusieurs moyennes</b> .....	148
<b>XI. Test de Wilcoxon-Mann-Whitney (WMW)</b> .....	150
<b>XII. Test de Wilcoxon pour séries appariées</b> .....	152
<b>XIII. Test de Kruskal-Wallis (KW)</b> .....	154
<b>XIV. Test de <math>\chi^2</math> de conformité ou d'ajustement</b> .....	156
<b>XV. Test de <math>\chi^2</math> d'homogénéité</b> .....	158

<b>XVI. Test de <math>\chi^2</math> à 4 cases pour comparer deux pourcentages</b> .....	160
<b>XVII. Test de <math>\chi^2</math> de McNemar pour séries appariées</b> .....	162
<b>XVIII. Test de <math>\chi^2</math> d'indépendance</b> .....	164
<b>XIX. Test de <math>\chi^2</math> de tendance</b> .....	166
<b>XX. Test du coefficient de corrélation</b> .....	168
<b>XXI. Test du coefficient de corrélation des rangs de Spearman</b> .....	170
<b>XXII. Test de la pente de la droite de régression</b> .....	172
<b>XXIII. Épreuve de normalité</b> .....	174
1. Méthode approchée.....	174
2. Méthode graphique .....	174
3. Méthode analytique.....	175

**CHAPITRE 14 – MODÈLES DE RÉGRESSION**

<b>I. Régression linéaire</b> .....	177
1. Régression linéaire simple : analyse de la variance .....	177
2. Régression linéaire multiple.....	184
<b>II. Régression logistique</b> .....	186
1. Régression logistique simple (univariée).....	186
2. Régression logistique multivariée.....	189
3. Régression logistique pour séries appariées .....	190
<b>III. Régression de Poisson (log-linéaire)</b> .....	190
1. Régression de Poisson simple (univariée).....	191
2. Régression de Poisson multivariée.....	194
3. Particularités de la régression de Poisson.....	195
4. Conditions d'application de la régression de Poisson.....	196
<b>IV. Modèle de Cox (risques instantanés proportionnels)</b> .....	196
<b>V. Autres modèles</b> .....	197
<b>VI. Méthodes d'analyse multivariée</b> .....	198
1. Codage des variables explicatives.....	198
2. Interactions .....	200
3. Sélection des variables à introduire dans un modèle.....	201
4. Paramètres d'analyse.....	202
5. Stratégies d'analyse.....	203
6. Principaux problèmes rencontrés dans une analyse multivariée .....	204

**PARTIE IV  
ÉPIDÉMIOLOGIE**

**INTRODUCTION** ..... 209

**CHAPITRE 15 – MESURES EN ÉPIDÉMIOLOGIE**

**I. Mesures de base** ..... 211

1. Proportion ..... 211

2. Ratio ..... 211

3. Cote ..... 212

4. Indice ..... 212

5. Taux ..... 212

**II. Indicateurs épidémiologiques** ..... 213

1. Prévalence ..... 213

2. Incidence ..... 214

3. Risque de maladie ..... 216

4. Relation entre incidence et prévalence ..... 217

5. Relation entre risque de maladie et densité d'incidence ..... 217

6. Mortalité globale ..... 218

7. Mortalité spécifique ..... 218

8. Risque de décès ..... 219

9. Mortalité proportionnelle ..... 219

10. Létalité ..... 220

**CHAPITRE 16 – ENQUÊTES ÉPIDÉMIOLOGIQUES**

**I. Protocole d'enquête** ..... 223

1. Plan d'analyse ..... 223

2. Questionnaire ..... 224

3. Définition des cas ..... 224

**II. Types d'enquêtes** ..... 226

1. Enquêtes descriptives ..... 226

2. Enquêtes étiologiques ..... 227

**III. Enquêtes de cohorte** ..... 228

1. Principe ..... 228

2. Présentation des données ..... 229

3. Mesures dans une enquête de cohorte ..... 229

4. Le risque relatif ..... 230

5. Choix du groupe de référence ..... 231

6. Nombre de sujets nécessaires à une enquête de cohorte ..... 231

7. Avantages et inconvénients d'une enquête de cohorte ..... 232

**IV. Enquêtes cas-témoins** ..... 233

1. Principe ..... 233

2. Présentation des données ..... 233

3. Mesures dans une enquête cas-témoins ..... 233

4. L'odds ratio (OR) ..... 234

5. Choix des témoins ..... 235

6. Nombre de témoins par cas ..... 236

7. Nombre de sujets nécessaires à une enquête cas-témoins ..... 236

8. Enquête cas-témoins appariée ..... 236

9. Avantages et inconvénients d'une enquête cas-témoins ..... 237

10. Variantes des enquêtes cas-témoins ..... 237

**V. Enquêtes transversales** ..... 238

**VI. Critères de causalité dans une enquête étiologique** ..... 239

**VII. Biais dans les enquêtes étiologiques** ..... 241

1. Biais de sélection ..... 241

2. Biais d'information ..... 241

3. Biais de mauvaise classification ..... 242

4. Biais de confusion ..... 242

**VIII. Prise en compte d'un tiers facteur: analyse stratifiée** ..... 242

1. Modificateur de l'effet ..... 242

2. Facteur de confusion ..... 243

3. Stratégie d'analyse stratifiée ..... 244

4. Analyse multivariée ..... 247

**CHAPITRE 17 – INVESTIGATION D'UNE ÉPIDÉMIE**

**I. Définitions** ..... 251

**II. Objectifs** ..... 253

**III. Chronologie** ..... 253

1. Affirmer la réalité de l'épidémie ..... 254

2. Confirmer le diagnostic ..... 254

3. Définir un cas ..... 255

4. Collecter les cas et les données ..... 255

5. Décrire l'épidémie ..... 256

6. Formuler des hypothèses ..... 259

7. Tester les hypothèses par une enquête étiologique ..... 259

8. Rechercher la preuve biologique ..... 260

9. Communiquer les conclusions de l'investigation ..... 261

10. Prendre les mesures de prévention ..... 261

**IV. Dynamique d'une épidémie** ..... 263

1. Phases clinico-biologiques d'une maladie infectieuse ..... 263

2. Ratio de reproduction ..... 264

3. Intervalle de génération ..... 267

4. Courbe épidémique en fonction de R et de IG ..... 267

5. Estimation du ratio de reproduction ..... 268

6. Mesures de lutte et de contrôle ..... 268

7. Modèles compartimentaux ..... 270

**CHAPITRE 18 – MESURES D'IMPACT**

**I. Fraction étiologique du risque** ..... 279

1. Fraction étiologique chez les exposés:  $FE_e$  ..... 279

2. Fraction étiologique dans la population  $FE_p$  ..... 280

<b>II. Fraction préventive</b> .....	281
1. Fraction préventive chez les sujets protégés $FP_e$ .....	281
2. Fraction préventive dans la population $FP_p$ .....	282
<b>III. Intervalle de confiance des FE et FP</b> .....	283

## CHAPITRE 19 – STANDARDISATION DES TAUX

<b>I. Position du problème</b> .....	285
<b>II. Principe</b> .....	285
<b>III. Méthode directe</b> .....	286
1. Calcul .....	286
2. Interprétation .....	287
<b>IV. Méthode indirecte</b> .....	287
<b>V. Conditions d'application</b> .....	288
<b>VI. Extension de la méthode</b> .....	288

## CHAPITRE 20 – ANALYSE DE SURVIE

<b>I. Principe</b> .....	291
<b>II. Méthode de Kaplan-Meier</b> .....	292
<b>III. La méthode actuarielle</b> .....	293
<b>IV. Comparaison de courbes de survie: test du log rank</b> .....	294

## CHAPITRE 21 – PERFORMANCES D'UNE TECHNIQUE

<b>I. Mesure expérimentale des performances d'un test</b> .....	297
1. Sensibilité .....	297
2. Spécificité .....	298
3. Cas d'un test quantitatif .....	299
4. Choix d'un seuil .....	300
<b>II. Performances d'un test en situation réelle</b> .....	302
1. Valeur prédictive positive .....	303
2. Valeur prédictive négative .....	303
3. Interprétation des VPP et VPN .....	304
<b>III. Reproductibilité et concordance</b> .....	307
1. Coefficient de concordance .....	307
2. Coefficient kappa .....	308
3. Coefficient kappa pondéré .....	309

## PARTIE V

## INITIATION À LA STATISTIQUE BAYÉSIENNE

### CHAPITRE 22 – LIMITES DES MÉTHODES FRÉQUENTISTES

<b>I. Estimation d'un paramètre</b> .....	315
<b>II. Tests statistiques</b> .....	315

### CHAPITRE 23 – EXEMPLES DE RAISONNEMENT BAYÉSIEN

<b>I. L'inversion bayésienne dans les tests d'hypothèses</b> .....	321
<b>II. Le théorème de Bayes</b> .....	321
<b>III. La vraisemblance</b> .....	322
<b>IV. Comparaison de deux hypothèses</b> .....	323
1. Le facteur de Bayes (K) .....	323
2. L'évidence .....	324
3. Interprétation d'un facteur de Bayes .....	324
4. Calcul de la probabilité <i>a posteriori</i> d'une hypothèse .....	325
<b>V. Comparaison de plusieurs hypothèses</b> .....	326
Conclusion .....	328

## ANNEXES

### RÉPONSES AUX QUESTIONS DES EXERCICES

.....	333
-------	-----

### RAPPELS MATHÉMATIQUES

Puissance .....	347
Racines .....	347
Logarithmes .....	347
Factorielles .....	348
Combinaisons .....	348
Probabilités .....	348
Formule de Bayes .....	348
Théorème de Bayes .....	349
Loi binomiale .....	349
Loi hypergéométrique .....	349

### FORMULAIRE STATISTIQUE

1. Loi binomiale .....	351
2. Fonctions de répartition de la loi binomiale .....	351
3. Loi de Poisson .....	351
4. Fonctions de répartition de la loi de Poisson .....	351
5. La loi normale ou loi de Gauss .....	352
6. La variable centrée réduite Z .....	352
7. Loi normale centrée réduite de Z .....	352
8. Paramètres de position et de dispersion d'une variable .....	352
9. Estimation d'un paramètre sur un échantillon .....	353
10. Écart type d'une moyenne et d'un pourcentage lorsque la taille n de l'échantillon est grande par rapport à la taille N de la population ( $n/N > 0,1$ ) .....	353
11. Intervalle de confiance d'une moyenne dans le cas des petits échantillons .....	353
12. Taille des échantillons pour réaliser un test Z ...	354
13. Corrélation .....	354
14. Régression .....	355

15. Intervalles de confiance d'un risque relatif et d'un odds ratio .....	355	26. Coefficient kappa pondéré entre 2 observateurs A et B.....	361
16. Nombre de sujets nécessaires à une enquête de cohorte .....	356	27. Formule de Welch-Satterthwaite.....	361
17. Nombre de sujets nécessaires à une enquête cas-témoins.....	356	28. Principales fonctions statistiques d'Excel®.....	362
18. Analyse d'enquête étiologique stratifiée.....	357	<b>BIBLIOGRAPHIE</b> .....	365
19. Valeur prédictives d'un test .....	357	<b>GLOSSAIRE</b> .....	367
20. Test exact de Fisher.....	358	<b>TABLES STATISTIQUES</b> .....	377
21. Correction de Yates: test du $\chi^2$ de Yates.....	358	Table 1: loi normale centrée réduite Z.....	377
22. Test du log rank pour comparer deux courbes de survie.....	358	Table 2: loi T de Student .....	378
23. Nombre de sujets nécessaires pour comparer deux pourcentages .....	359	Tables 3: loi F de Fisher (test unilatéral).....	380
24. Coefficient de dissymétrie ( <i>skewness</i> ).....	359	Table 4: loi du $\chi^2$ .....	384
25. Coefficient d'aplatissement ( <i>kurtosis</i> ).....	360	<b>INDEX</b> .....	385