

« Entre un penseur et un érudit, il y a la même différence qu'entre un livre et une table des matières. » Jean-Baptiste Say

Table des matières

1	Avant-propos	1
1.1	Variabilité et erreur	1
1.2.	Enseignement et apprentissage	3
1.3.	Auto-évaluation formative	4
1.4.	Style du texte	5
1.5.	Navigation dans cet ouvrage	5
1.6.	Tables et formulaires	8
1.7.	Bibliographie	8
10	Variabilités et erreurs	11
10.1	Population et mesures expérimentales	11
10.2	En quoi consiste la variabilité ?	13
10.2.1	Imprécision	14
10.2.2	Inexactitude	15
10.2.3	Variabilité factorielle	16
10.2.4	Variabilité individuelle	17
10.3	Comment maîtriser la variabilité ?	19
10.3.1	Pourquoi la mesure est-elle variable ?	19
10.3.2	En quoi la variabilité pose-t-elle un problème ?	20
10.3.3	En quoi les statistiques apportent-elles une solution?	21
10.3.4	Peut-on supprimer la variabilité ?	21
10.4	La force de la preuve	23
10.5	Mesures de la variabilité	24
10.5.1	Écarts à la moyenne arithmétique	24
10.5.2	Mesure de la variabilité	25
10.5.3	Equation des sommes des carrés des écarts	27
10.5.4	Variance	30
10.5.5	Ecart-type	31
10.5.6	Centrage	32
10.5.7	Standardisation	33
11	Description des données	35
11.1	Définitions	35
11.2	Description d'une variable	35
11.2.1	Tables et graphiques des variables discrètes	35
11.2.2	Tables et graphiques des variables continues	40
11.2.3	Tendance centrale d'une distribution	44
11.2.4	Dispersion d'une distribution	46
11.2.5	Représentation graphique de la moyenne et de l'écart-type	49
11.3	Risques d'erreur	51
11.4	A voir sur le site	52

20	Variabilité conjointe de deux variables	53
20.1	Tables et graphiques	54
20.2	Intensité de la relation entre deux variables	56
20.2.1	Modélisation de la variabilité	56
20.2.2	Mesure de la qualité du modèle	59
20.2.3	Propriétés du coefficient de détermination	61
20.3	Mesure de l'intensité de la relation entre deux variables	62
20.3.1	Somme des produits d'écart à la moyenne	62
20.3.2	Covariance	64
20.3.3	Coefficient de corrélation linéaire	66
20.3.4	Interprétation de la valeur du coefficient de corrélation	67
20.3.5	Relations non linéaires	68
20.3.6	Corrélation et causalité	70
20.4	Caractérisation de la relation entre deux variables	71
20.4.1	Ecart résiduel	71
20.4.2	Droite des moindres carrés $Y(X)$	73
20.4.3	Interpolation et extrapolation	75
20.5	Risques d'erreur	75
20.5.1	Confusion entre corrélation et détermination	75
20.5.2	Influence des données extrêmes	76
20.5.3	Extrapolation abusive	76
20.5.4	Confusion entre corrélation et causalité	76
20.6	A voir sur le site	76
30	Etude du hasard	77
30.1	Déduction et induction	77
30.2	Notions d'épreuve et d'événement	79
30.3	Notion de probabilité	81
30.3.1	La probabilité définie à partir de l'observation	81
30.3.2	La probabilité définie a priori	82
30.4	Lois de probabilités	84
30.4.1	Loi des probabilités composées (événement indépendants)	84
30.4.2	Tables de contingence	85
30.4.3	Probabilités conditionnelles	86
30.4.4	(Non) Indépendance	87
30.4.5	Loi des probabilités composées, forme générale	88
30.4.6	Loi des probabilités totales	89
30.4.7	Eprouver un modèle	91
30.4.8	Loi de Bayes	93
30.5	Faux négatif et faux positif	93
30.6	Risque d'erreur	97
30.7	A voir sur le site	97

40	Modélisation de la variabilité	99
40.1	Introduction	100
40.2	Modèles concernant des dénombrements	101
40.2.1	Modélisation	101
40.2.2	Tables et graphiques	102
40.2.3	Valeurs caractéristiques	103
40.2.4	Modèles détaillés	103
40.3	Modèles concernant des mesures continues	104
40.3.1	Fonction de densité de probabilité	104
40.3.2	Tables et graphiques	105
40.3.3	Valeurs caractéristiques	106
40.3.4	Modèles présentés	106
50	Le modèle binomial	107
50.1	Principe	107
50.2	Exemple de l'analyse de portions du génome	109
50.2.1	Hypothèses du modèle	109
50.2.2	Probabilités découlant du modèle	109
50.2.3	Intérêt du modèle	110
50.3	Tables	110
50.4	Graphiques	111
50.5	Risque d'erreur	112
60	Le modèle de Poisson	113
60.1	Principe	113
60.2	Tables et graphiques	114
60.3	Exemple	115
60.4	Risque d'erreur	115
70	Le modèle Normal de Gauss et Laplace	117
70.1	Fonction de probabilité	117
70.2	Valeurs caractéristiques	118
70.3	Calcul de probabilité	119
70.4	Variable normale réduite	120
70.4.1	Utilité	120
70.4.2	Standardisation (ou réduction)	120
70.4.3	Propriétés	121
70.4.4	Tables et graphiques	122
70.5	Faux positifs et faux négatifs	125
70.6	Théorème central limite	125
70.7	Distribution d'échantillonnage de M_x	128
70.7.1	Utilité	128
70.7.2	Augmentation de la précision	128
70.8	A voir sur le site	130

80	Le modèle de Student	131
80.1	Principe	131
80.2	Tables et graphiques	132
80.3	Exemple	133
90	Le modèle du Khi-Carré	135
90.1	Principe	135
90.2	Degrés de liberté	137
90.3	Tables	137
90.4	Graphiques	138
90.5	Exemples	139
100	Le modèle de Fisher	141
100.1	Principe	141
100.2	Degrés de liberté	142
100.3	Tables	143
100.4	Graphiques	144
100.5	Exemple	144
105	Imprécision et incertitude	145
105.1	Principe de l'intervalle de confiance	145
105.2	Distribution d'échantillonnage	146
105.3	L'intervalle de confiance	147
105.4	Intervalle de confiance de la moyenne	148
105.4.1	Principe	148
105.4.2	Technique	149
105.4.3	L'intervalle de confiance peut réserver des surprises	152
110	Décisions et risques d'erreur	153
110.1	Prise de décision par un test d'hypothèse	153
110.2	Risque d'erreur connu et inconnu	154
120	Maitrise des risques d'erreur	159
120.1	Complexification des modèles décrivant le hasard	159
120.1.1	Modèle binomial	159
120.1.2	Modèle Khi-Carré	163
120.1.3	Modèle de la distribution d'échantillonnage des moyennes	166
120.2	Optimisation du plan expérimental	169
120.2.1	Taille de l'échantillon	169
120.2.2	Interdépendance des paramètres	170
120.2.3	Contrôle de la puissance	173
120.2.4	Optimisation de la puissance	177
120.2.5	La plus petite différence pertinente	177
120.2.6	Information de l'intervalle de confiance suite à AHO ©	178
120.2.7	Déterminer n	179
120.2.8	A voir sur le site	181
120.2.9	Remarque sur le mode de calcul	181

125	Tests d'hypothèse relatifs aux variances	183
125.1	Calcul de la variance	183
125.2	Distributions de référence	183
125.3	Comparaison de deux variances	184
125.3.1	Modèle	184
125.3.2	Test bi-directionnel	186
125.3.3	Conclusions et risque d'erreur	186
125.4	Comparaison de plus de deux variances	186
125.4.1	Contexte	186
125.4.2	Modèle	187
125.4.3	Conclusions et risque d'erreur	188
130	Tests d'hypothèse relatifs aux moyennes	189
130.1	Conditions générales	189
130.1.1	Echantillonnage	189
130.1.2	Distributions	190
130.1.3	Risques d'erreur	190
130.2	Test de comparaison d'une moyenne à un standard	191
130.2.1	Principes	191
130.2.2	Un échantillon de valeurs indépendantes, variance connue	193
130.2.3	Egalité des variances	195
130.2.4	Un échantillon de valeurs indépendantes, variance estimée	196
130.2.5	Deux échantillons de valeurs paires	198
130.3	Test de comparaison de deux moyennes	199
130.3.1	Le test de Student	199
130.4	Risques d'erreur dans le choix du test	202
135	Tests d'hypothèse relatifs aux fréquences	203
135.1	Modèles générant des fréquences théoriques	203
135.2	Modèles décrivant l'effet du hasard	204
135.2.1	Une ligne et deux catégories	204
135.2.2	Une ligne et plus de deux catégories	205
135.2.3	Table de contingence 2 x 2 (indépendance)	206
135.2.4	Table de contingence 2 x 2 (homogénéité de proportion)	208
135.2.5	Table de contingence plus de 2 x 2	209
135.2.6	Remarque sur le mode de calcul	210
135.3	Risque d'erreur	211

140	Comparaisons de moyennes : ANOVA I fixe	213
140.1	Concept de l'ANOVA	213
140.2	Modèle	215
140.3	Hypothèses	216
140.4	Equation de l'ANOVA	217
140.5	Carrés moyens	217
140.6	Exemple d'ANOVA1 : expérience américaine	218
140.6.1	Test d'homogénéité des variances	218
140.6.2	Test d'égalité des moyennes	219
140.7	Exemple d'ANOVA1 : expérience européenne	219
140.7.1	Test d'homogénéité des variances	219
140.7.2	Test d'égalité des moyennes	220
140.8	Que décider face à des conclusions contradictoires ?	220
140.9	Risques d'erreur	221
140.10	A voir sur le site	221
150	Contrastes ou comparaisons post-hoc	223
150.1	Principe	223
150.2	Contrastes de Scheffé	224
150.3	Risque d'erreur	225
150.4	A voir sur le site	225

Remerciements

Les personnes impliquées de près ou de loin, dans un ouvrage qui repose sur toute une carrière sont si nombreuses qu'il est impossible de toutes les citer : collègues et relations internationales marquantes, j'en ai cité plusieurs en tête de cet ouvrage, en craignant plus que tout d'oublier un nom. Qu'il me pardonne si c'était le cas.

Je m'en tiendrai ici aux personnes qui ont directement contribué à la réalisation de cet ouvrage : Véronique Geubelle (graphiste, design de la couverture), Elise Depiereux (relecture orthographique), Annick Genette (relecture statistique), Laurence Meurant (linguiste), Natacha Lerot (Impression et édition), Stéphanie Herfurth (promotion et diffusion). J'ai bénéficié des remarques pertinentes de Benoit Bihin, Martine De Vleeschouwer-Dieudonné, Jean-Michel Vandeweerd ainsi que de celles des créateurs et développeurs du site web : Philippe Calmant, Grégoire Vincke et Anne-Cécile Wauthy. Philippe Calmant a initié le site et Nicolas Sauveur y a développé plusieurs des animations Flash.

Je porte une attention particulière et très reconnaissante aux centaines de courriels constructifs reçus de la part de mes étudiants.