

INDEX

A

added variable plot, 462
AIC, 465, 470
analyse bivariée, 36
analyse des résidus, 457, 464
analyse multivariées, 36
analyse multivariée, 36
analyse univariée, 36

B

backward selection, 471
bande de confiance, 413
bar chart, 67
best subset selection, 472
biais, 210
biais de confusion, 426, 428
BIC, 465, 470
bootstrap, 180
boxplot, 70

C

centiles, 49
centre de gravité, 399
centroïde, 401
ceteris paribus, 427
coefficient de corrélation partielle de Spearman, 360
coefficient d'aplatissement, 59
coefficient d'aplatissement d'échantillon, 162
coefficient d'asymétrie, 59
coefficient d'asymétrie d'échantillon, 161
coefficient de corrélation de rang, 356
coefficient de corrélation de Spearman, 356
coefficient de corrélation linéaire, 62, 127, 342
coefficient de corrélation linéaire d'échantillon, 160
coefficient de corrélation linéaire de Pearson, 342
coefficient de corrélation linéaire de population, 161
coefficient de corrélation multiple, 351
coefficient de détermination, 346, 351, 397, 456
coefficient de détermination ajusté, 456
coefficient de détermination partiel, 352, 353
coefficient de dispersion, 417
coefficient de kurtosis de population, 162
coefficient de skewness de population, 162
coefficient de variation, 417
coefficient phi, 368
coefficient Tau de Kendall, 364
coefficients de régression, 378
comparaisons multiples, 333
condition d'exogénéité, 433
conservatrice, 336
convergent, 211
convergent en moyenne quadratique, 211, 212
convergent en probabilité, 212
corrélation bisériale ponctuelle, 369
corrélation intra-classe, 289
corrélation partielle, 352, 353

covariance, 61, 126
covariance d'échantillon, 160
covariance de population, 160
critère d'Akaike, 465, 470
critère de Schwarz, 465, 470

D

DAG, 465
déciles, 49
degrés de liberté, 395
degrés de liberté de Satterthwaite, 279
dénombrable, 82
densité, 110
déviatoin-standard, 52
dfbeta_i, 407
dfits_i, 406
diagramme à bâtons, 67
diagramme à points, 67
diagramme à tige et feuilles, 69
diagramme boîte à moustaches, 70
diagramme de Venn, 82
diagramme des fréquences cumulées, 71
diagramme en secteurs, 69
diagramme nuage de points, 72
différence de risque, 236
distribution, 33, 35
distribution bimodale, 39
distribution d'échantillonnage, 163
distribution d'échantillonnage empirique, 163
distribution d'échantillonnage théorique, 163, 166
distribution de probabilité, 106
distribution de probabilité jointe, 128
distribution discrète, 183
distribution exacte, 166
distribution multimodale, 39
distribution uniforme, 183
distribution unimodale, 39
dominance stochastique, 319
domine stochastiquement, 318
données, 30
dotplot, 67
droite des moindres carrés, 390

E

écart absolu moyen, 51
écart-type, 52, 124
échantillon, 30
échantillon aléatoire, 155
échantillon aléatoire simple, 137
échantillons appariés, 229, 238
efficace, 212
élasticité, 451
élément, 81
endogène, 384
ensemble, 81
ensemble fondamental, 83
ensemble vide, 81
épreuve Binomiale, 184

épreuve Binomiale négative, 186
 épreuve de Poisson, 189
 épreuve hypergéométrique, 187
 épreuve Multinomiale, 188
 épreuves de Bernoulli, 184
 erreur, 379
 erreur de première espèce, 255
 erreur de seconde espèce, 257
 erreur de type I, 255
 erreur de type II, 257
 erreur quadratique moyenne, 211
 erreurs de mesure, 433
 erreurs de spécification, 433
 espace échantillon, 83
 espérance, 118
 espérance du taux de faux positif, 337
 espérance mathématique, 118
 estimateur, 204
 estimateur ponctuel, 205
 estimation, 204
 estimation ponctuelle, 412
 étendue, 48
 étendue interquartile, 50
 étude causale, 465
 étude descriptive, 464, 465
 étude étiologique, 465
 étude prédictive, 465
 événement, 83
 événement complémentaire, 84
 événement composé, 83
 événement élémentaire, 83
 exogène au sens faible, 384
 exogénéité, 383
 exogénéité stricte, 386
 expérience aléatoire, 83

F

facteur confondant, 425
 facteur intermédiaire, 466
 facteur pronostique, 425
 famille d'hypothèses nulles, 333
 FDR, 337
 fluctuations d'échantillonnage, 205
 fonction cumulative, 107, 115
 fonction de densité, 110
 fonction de puissance, 291
 fonction de quantile, 156
 fonction de régression, 377
 fonction de régression linéaire, 378
 fonction de répartition, 107, 115
 forme fonctionnelle, 387
 forward selection, 471
 fréquence, 34
 fréquence relative, 34
 FWER, 333

G

graphe acyclique, 465

H

hat matrix, 400
 hétéroscédasticité, 388
 histogramme, 33, 68
 homoscedasticité, 387, 434
 hyperplan, 430
 hypersurface, 431
 hypothèse alternative, 253
 hypothèse composite, 254
 hypothèse nulle, 253
 hypothèse paramétrique, 254
 hypothèse simple, 254
 hypothèse statistique, 253
 hypothèses du modèle de régression linéaire multiple, 435
 hypothèses du modèle de régression linéaire simple, 382

I

iid, 137
 indépendance, 93, 129, 130
 indépendant, 94
 inégalité de Bienaymé-Tchebychev, 141
 influence, 405
 interaction, 441
 interaction statistique, 441
 interprétation de la constante, 439
 intersection, 82
 intervalle de confiance, 205, 213, 397, 455
 intervalle de confiance d'un Odds Ratio, 245
 intervalle de confiance d'une corrélation, 349
 intervalle de confiance d'une moyenne, 216
 intervalle de confiance d'une proportion, 233
 intervalle de confiance de la différence de deux moyennes, 223
 intervalle de confiance de la différence de deux proportions, 236
 intervalle de confiance du ratio de deux proportions, 240
 intervalle de confiance pour la droite de régression, 412
 intervalle de prédiction, 414
 intervalle interquartile de population, 159

K

Kappa, 289
 kurtosis, 59
 kurtosis d'échantillon, 162

L

levier, 399
 loi Binomiale, 184
 loi Binomiale négative, 186
 loi de Bernoulli, 185
 loi de Cauchy, 150, 197
 loi de Fisher, 197
 loi de Fisher-Snédecor, 197
 loi de Pascal, 186
 loi de Poisson, 190
 loi de Student, 196
 loi du Chi², 194
 loi faible des grands nombres, 90, 142

loi Gamma, 199
loi Géométrique, 186
loi Hypergéométrique, 188
loi Multinomiale, 189, 307, 312, 315
loi Normale, 192
loi Normale bivariée, 342, 345
loi Normale centrée et réduite, 113
loi Normale standardisée, 350
loi Normale trivariée, 350
loi uniforme, 183, 191
lois des grand nombres, 90

M

McNemar, 287
médiane, 41
médiane de population, 156
médiateur, 466
médiation, 466
mesure d'association entre une variable continue et une variable binaire, 369, 371
mesure de dispersion, 46, 66
mesure de forme, 59
mesure de position, 55
mesure de tendance centrale, 38, 66
mesures d'association entre variables discrètes
mesures d'association entre variables discrètes ordinales, 363
mesures d'association entre variables qualitatives nominales, 367
mesures de dépendance, 61
mesures de forme, 66
méthode d'Agresti et Caffo, 238
méthode d'Agresti et Coull, 235
méthode de Benjamini et Hochberg, 337
méthode de Benjamini et Yekutieli, 338
méthode de Bonferroni, 335
méthode de Clopper-Pearson, 235
méthode de Holm, 336
méthode de Wald, 233, 237, 242
méthode de Wald, 239, 243, 245, 248
méthode de Wald avec correction de continuité, 233, 237
méthode de Wald avec correction de continuité, 239
méthode de Wilson, 234
méthode de Wilson avec correction de continuité, 234
méthode de Woolf, 248
méthode des moindres carrés, 389, 437
méthode des moments, 209
méthode du maximum de vraisemblance, 207
méthode du score, 243
mode, 38, 157
modèle, 153
modèle additif simple, 438
modèle ANOVA, 479
modèle d'analyse de variance, 479
modèle de Bernoulli, 168
modèle de distribution, 112, 153
modèle de prédiction, 414
modèle de régression, 380
modèle descriptif, 464
modèle log-log, 451
modèle Normal, 153

modèle Normal bivarié, 344
modèle prédictif, 465
modèle probabiliste, 153
modèle semi-logarithmique, 450
moment centré d'ordre r , 162
moment croisé, 126
moment d'ordre 1, 122
moment d'ordre 2, 122
moyenne arithmétique, 39, 138
moyenne d'échantillon, 156
moyenne de population, 156
moyenne empirique, 90
moyenne pondérée, 41
moyenne théorique, 90
multicolinéarité, 435
mutuellement exclusifs, 84

N

niveau de confiance, 205
nombre de sujets pour tester la différence de deux moyennes, 300
nombre de sujets pour tester la différence de deux proportions, 303
nombre de sujets pour tester une moyenne, 296
nombre de sujets pour tester une proportion, 298
nombre de sujets pour une précision donnée de la différence de deux moyennes, 301
nombre de sujets pour une précision donnée de la différence de deux proportions, 304
nombre de sujets pour une précision donnée pour une moyenne, 297
nombre de sujets pour une précision donnée pour une proportion, 299
non dénombrable, 82

O

observation, 30
Odds Ratio, 245
outlier en X, 400
outlier en Y, 400
outliers, 221

P

paire d'observations, 230
paramètre, 152
paramètres d'échelle, 159
partial regression plot, 462
percentile α , 55
percentiles, 48
performance prédictive, 96
PEV, 472
pie chart, 69
plan de régression, 431
polygone des fréquences, 35
polynôme fractionnel, 448
population, 152
population, 29
PPV, 96
prédiction, 390
probabilité, 87

probabilité à postériori, 101
probabilité à priori, 101
probabilité conditionnelle, 93
probabilité de faux positif, 103
probabilité jointe, 128
probabilité marginale, 128
procédure de test, 253
proportion, 168
proportion de variation expliquée, 472
puissance, 290
puissance d'un test, 289
p-valeur, 261
p-valeur exacte, 329

Q

Q-Q plot, 411
quantile d'ordre α , 55
quantile α , 55
quantiles, 48
quartiles, 48
quintiles, 49

R

rapport de vraisemblance négatif, 98
rapport de vraisemblance positif, 98
région critique, 254
région de confiance simultanée, 413
règle d'additivité, 92
règle de Bayes, 101
règle de multiplication, 92
règle de soustraction, 91
régression, 377, 430
régression linéaire, 378
régression linéaire multiple, 423
régression vers la moyenne, 374
relation ajustée, 423
relation brute, 423
relation causale, 424
relation de régression partielle, 462
relation dose-réponse, 448
relation fonctionnelle quadratique, 448
relation non-ajustée, 423
résidu, 390
résidu jackknife, 404
résidu jackknife studentisé, 405
résidu studentisé externe, 405
résidu studentisé interne, 404
résidus théoriques, 403
résumés quantitatifs, 37
risque d'erreur de première espèce, 255
risque d'erreur de première espèce de la famille, 333
risque de deuxième espèce, 257
risque de type II, 257
risque relatif, 240
RVN, 98
RVP, 98

S

sans biais, 210
scatter plot, 72

score de propensité, 469
score standardisé, 56
semi-élasticité, 450
sensibilité, 96
seuil critique, 254
seuil de la région critique, 254
skewness, 59
skewness d'échantillon, 161
sous-ensemble, 81
spécificité, 96
standardisation, 135
statistique, 152, 203
statistique de Kruskal-Wallis, 326
statistique de test, 213
statistique de Wald, 453
statistique de Wilcoxon, 322
statistique descriptive, 29
statistique Kolmogorov-Smirnov, 331
statistiques de Mann & Whitney, 322
stem-and-leaf plot, 69
stochastiques, 431
stratifiant, 429
surface de régression, 431

T

table de contingence, 77
tableau, 34
taille d'échantillon (voir aussi "nombre de sujets")
taille d'échantillon pour détecter une corrélation linéaire non nulle, 349
taille d'échantillon pour une précision donnée pour une corrélation, 350
Tau (τ) de Kendall, 363
taux d'erreur de première espèce, 255
taux de couverture, 205
tertiles, 49
test bilatéral, 260
test d'homogénéité, 310, 312
test d'indépendance, 314, 316
test d'une moyenne, 269
test d'une proportion, 282
test de Breusch-Pagan, 411
test de contingence, 315
test de contingence de Pearson, 367
test de goodness of fit, 307, 309
test de Kolmogorov-Smirnov, 331
test de Kruskal-Wallis, 325
test de la corrélation de rang, 360
test de la corrélation linéaire, 346
test de la différence de deux moyennes, 276
test de la différence de deux proportions, 284
test de la somme des rangs, 318
test de McNemar, 288
test de Pearson, 315
test de significativité, 395, 452
test de Student, 395
test de Welch, 279
test de Wilcoxon-Mann-Whitney, 318
test diagnostique, 95
test exact de Fisher, 328
test F, 453
test unilatéral, 260

tests nonparamétriques, 317
théorème central limite, 142
théorème de Frisch-Waugh, 463
théorème de Khintchine, 142
théorème de Lindberg-Feller, 144
tirage aléatoire simple, 155
transformation, 416
transformation logarithmique, 417
transformation quadratique, 418
transformation racine carrée, 417
transformation réciproque, 417
transformation Z de Fisher, 348

U

unité statistique, 29
univers, 82

V

V de Cramér, 367
valeur prédictive négative, 96
valeur prédictive positive, 96
validation croisée, 465
variable, 29

variable aléatoire, 104, 105
variable aléatoire continue, 110
variable aléatoire discrète, 106
variable binaire, 167
variable continue, 30
variable de Bernoulli, 168
variable dépendante, 377, 431
variable discrète, 31
variable discrète polytomique, 443
variable indicatrice, 440
variable pivot, 213
variable qualitative, 30
variable quantitative, 30
variables explicatives, 431
variables indépendantes, 377
variance, 51, 122
variance d'échantillon, 158
variance de population, 158
VIF, 436
vraisemblance, 207

Z

Z-score, 56