Index

a

abduction 326
accreditation 384–385
accuracy 105, 106, 284, 380, 382
  of analysis 106
adaptive neuro-fuzzy inference system (ANFIS) 361–362
adaptive splines 85
adjacency matrix 304, 352
aggregation, performance characteristics 108–110
alignment methods
  DTW 91
  left panel, simulated chromatogram 91
allele 366
alternating conditional expectations (ACE) 281, 286–290
alternating least squares (ALS) 182, 184
alternative, compound propositions 325
alternative hypothesis. *See* hypothesis
American Food and Drug Administration (FDA) 385
analog-to-digital converter (ADC) 4, 5, 8
analysis of variance (ANOVA)
  ASCA 52–56
  definition 45
  MANOVA. *See* multidimensional analysis of variance (MANOVA)
  multi dimensional 51
  multi-way 48, 51. *See also* multi-way analysis of variance
  one-way 45. *See also* one-way analysis of variance
  regression analysis 236
  sums of squares (SS) 235
  two-way 48. *See also* two-way analysis of variance
analytical databases
  coding of chemical structures. *See* chemical structures coding
  coding of spectra 298–301
  examples 295, 296
  exchange files 297–298
  information types 296
  laboratory information and management systems (LIMS) 308–309
  source and library files 297
  structure 296, 297
  working and index file 308
analytical range 105
analyzers 10, 323
Anthranilic acid 289–290
apodization function 82
area determination 61, 74, 75
artificial intelligence (AI) 2, 11–13, 316, 321
  definition 321
  knowledge acquisition and processing 329, 330
  methods 323
  modules 321, 322
artificial neural networks 321, 330–332, 343, 345. *See also* neural networks
artificial neuron 331
assembler language 8
associative learning law 337
atomic absorption spectrometry 37, 104

Chemometrics: Statistics and Computer Application in Analytical Chemistry, Fourth Edition. Matthias Otto.
© 2024 WILEY-VCH GmbH. Published 2024 by WILEY-VCH GmbH.

atomic emission spectroscopy 91, 323
  inductively coupled plasma (ICP-AES) 91
autocorrelation 61, 92, 93, 95, 99
  correlated data 97
  correlogram 95
  drift and periodicities 97
  empirical 96
  function 96, 98
  measure 93
  point-wise 94
  stationary process 94
  time lag 93, 95
  uncorrelated data 96
autocovariance 93–95
automation 1, 3, 9
autoregression 99
autoscaling 148, 149
axiom 325
axon 331

b

background correction 61, 76, 82, 84
backpropagation 337, 343, 344, 370
  networks 341
backtracking 326
bagging 218–220
BASIC 8, 326
Bayes' theorem 205
Bayesian classification 205
Bayesian regulation backpropagation algorithm 344
Beer's law 259, 260, 263, 264, 281
best linear unbiased estimator (BLUE) 250
bias 16, 45, 106, 112, 274, 294, 337, 343. *See also* error; systematic
bidirectional associative memory (BAM) 333, 337, 377
binary number system 4
biological fluids 91
biplot 159, 160
bit 1, 4, 6–8, 309, 312–313, 366–367, 369, 370, 372
blank 104, 105, 361, 380–382
  samples 381
blind source separation problem 178
blocking 113, 118, 119, 134
blood 10, 16, 91, 188, 189, 198, 217, 384
bond electron (BE) matrix 304
boosting 219
bootstrapping 219, 285, 347
box-and-whisker plots 25
Box–Behnken design 125
bulk storage devices 7
bus 7
byte 6

c

C 8, 326
calcium 188
calibration 231
  curved 104, 105
  direct 261, 263
  function 232, 380
  indirect 263, 264
  inverse 265–268, 272
  K-matrix approach 264
  P-matrix approach 265
calibration function 104, 105
canonization 307
cardinality 359, 360, 377
centering 147, 149, 227, 255
central composite design 125–127
central limit theorem 21
central processing unit (CPU) 7
certified reference materials 384
ceteris paribus principle 111
change and hope optimization procedure 135
chaos 13
chemical databases 295. *See also* analytical databases
chemical structure 12, 295–298, 300–311, 314–316
  matrix representation 161, 303–306
chemical structures coding
  canonization 307
  connection table 306–307
  fragmentation codes 301–303
  Markush structures 307
  matrix representation 303–306
chemometrics
  in Analytical Chemistry 1
  computer-based methods. *See* computer-based laboratory

data evaluation and interpretation methods 11
definition 1
Chernoff faces 196, 197
chromatographic response function (CRF) 109
chromatography 10, 108, 109, 128, 145, 173, 177, 179, 279, 300, 323, 364
chromosome 366, 367, 370, 378
city-block distance 186, 187
class 12, 17, 23, 39, 145–147, 198, 200, 203, 205–208, 211–215, 217, 219, 226, 307, 312, 336, 338, 339, 347, 363
classification
by *k*-nearest neighbors (*k*NN) 192
by discriminant analysis 201, 213, 215, 290
by *k*-nearest neighbors (kNN) 207
by soft independent modeling of class analogies (SIMCA) 198, 210, 211
classification and regression trees (CART) 216–221, 225, 226, 229, 291
class width 16, 17
cluster analysis 12, 145, 185, 216, 311, 312, 336, 364
distance measures 186
hierarchical 188–193
nonhierarchical 194–196
similarity measures 187
clustering 146
crisp 194, 313, 353, 354
fuzzy 194, 364, 365
hierarchical 188
nonhierarchical 194, 195
c-means algorithm 194, 195
c-means clustering 196
coding
chemical structures 301–319
spectra 295
coefficient
of determination 239, 272
of variation 23
Coiflet 86
collinearity 264
common factors 153, 161–164, 171, 172, 174
communalities 163
competitive learning law(s) 346
steps 337
supervised 338
unsupervised 338
compiler 8
complement 358
compositional rule of inference 360
computer application in analytics (COMPANA) 1
computer-based analytical chemistry (COBAC) 1
computer-based information systems 12
computer-based laboratory 3
analog and digital data 3
analog-to-digital converters (ADCs) 5
binary *vs*. decimal number system 4
central processing units and buses 7
computer terminology 6
digital-to-analog converters (DACs) 6
input/output-systems 8
memory 7
networking 8
programming 8
robotics and automation 9
computer program 3, 6, 8, 328
conditioning 10, 250
condition number 227, 250, 251
confidence
band 243, 244, 285
interval 24, 241, 252, 294, 347
confidence interval 24, 381
confounding 114, 125
parameter estimations 125
confusion matrix 202, 352
conjunction 325
connection table 306–307, 314, 315, 318
bond atoms and bonds 306
connection matrix of atoms 306
nonredundant 307
continuous/discrete continuous quantities 104
control
chart 382–383
sample 382
convolution 61, 76, 80
Cook's distance 273, 275
correlation
coefficient 91, 93, 94, 150, 164, 239, 311

correlation (*contd.*)
matrix 150, 151, 156, 158, 160, 162–166, 168, 169, 171, 189, 333, 335
reduced 163, 164
measure 16, 159
corticoids 311
cortisone 311
counting rates 20
covariance matrix 69, 92, 150, 163, 202, 206, 207, 210, 241, 243, 285, 290
cross-correlation 61, 92, 98
crossover 367, 368
cross-validation 86, 156, 179, 208, 218–220, 225, 255, 257, 268, 273, 274, 276, 291
cross-validation principal components 211

d

data
analog 3
correlated 16, 61, 97, 99
digital 3, 5
interpretation 11
missing 147
preprocessing 145, 147, 149
random 11, 16, 61, 208
set characterization 21, 148, 151, 156, 157, 179, 185, 188, 203, 219, 225, 229, 270, 287–289, 300, 309, 313, 343, 362, 364
three-way 179, 183, 185, 279, 280
transfer
parallel 7
serial 7
two-way 48, 179, 278
database
analytical 295–320
model 308
structure 296
data preprocessing
correlation matrix 150, 151
mean centering 148
missing data 147
scaling 148
variance–covariance matrix 150
data reduction 79, 82, 84, 300, 308
Daubechies 86
decimal number system 4
decision boundary 199, 201, 214, 345
deconvolution 61, 76, 81, 82
deduction 325
degrees of freedom 24, 31, 32, 36, 38, 47, 51, 116, 235, 239, 270, 273, 383
delta rule 337, 342
dendrite 331
dendrogram 191, 193
density function. *See* frequency function
derivation algorithm 11
descriptive statistics
box-and-whisker plots 25
confidence interval 24
dispersion measure 23
error propagation 25, 27
geometric mean 22
harmonic mean 22
median 23
quartiles 23
random numbers. *See* random number distribution
uncertainty and error 28
detection limit 105, 361, 380
diagnostic statistics
hat matrix 268
influential observations 270, 271, 275
inverse calibration model 268
outliers 269, 270
residuals and prediction error 269
SEC 268
digit
least significant 4
most significant 4, 83
digital smoothing and filtering
advantages 62
filter selection process 64, 67
Kalman filter 68, 69
moving average filter 62
Savitzky–Golay filter. *See* Savitzky–Golay filter
Simpson rules 75
trapezoidal rules 74
digital-to-analog converter (DAC) 6, 8
discrete analyzers 10
discrete membership functions 357
discrete quantities 104
discriminant analysis

linear (LDA) 201, 225, 226
quadratic (QDA) 207, 225, 226
regularized (RDA) 207
dispersion 15, 18, 23, 106, 152, 162, 166, 381, 382
matrix 152, 162, 166. *See also* variance
measure 23
dissimilarity measures 312
distance
matrix 188–190, 192
measure 186, 189, 191, 311, 312, 337, 346
distribution
binomial 21
F 36, 37
function 19, 29, 40, 354
Gaussian. *See* normal distribution
kurtosis 20, 21, 24, 172
moment 19
Poisson 20
skewness 19
Student's 24, 31, 32, 381
Dixon's Q-test 43
drift 61, 69, 80, 86, 93, 97, 112
dynamical system model 68
dynamic range 105
dynamic time warping (DTW) 91

e

eigenvalue 52, 156, 157, 251, 265, 268
one criterion 156
eigenvector
analysis 152, 165
projection 151
elemental composition 51, 146, 147, 158, 173, 177, 204
ensemble methods 218, 219
enzyme 114, 120, 123, 131, 133, 140
ceruloplasmin (CP) 130, 131
equivalence 146, 325
error
first and second kind 41
integral. *See* distribution function
propagation 15, 25, 27, 72, 74, 108
random 16, 37, 45, 48, 99, 106, 245, 248
real 171
round-off 3, 26, 62, 251
of second kind 41
systematic 16, 105, 112, 379
α error 41
β error 41–42
ESPRIT 330
Euclidean distance 186–188, 193, 337
European Union regulation 33
evolutional component 365
evolutionary strategies 12, 13
evolving factor analysis (EVA) 177
exchange file 297–298
exclusive OR 312, 313, 339, 340
experimental design
balanced 114
blocking factors 113
Box–Behnken 125, 126, 128, 130
central composite 125–127
ceteris paribus principle 111
confounding 114
factorial experiments 113
fractional factorial 115, 118
full factorial 114, 115, 121, 124–126, 135
Latin square 118, 119
lattice 129
mixture 128, 129
Plackett and Burman 118
randomization 112
replication 111
response surface. *See* three-level
saturated 118
screening. *See* two-level
star 125, 126
supersaturated 119
symmetry 114
three-level 124, 126
three-level designs. *See* response surface designs
two-level 114, 115, 117
two-level designs. *See* screening designs
expert systems 5, 8, 12, 323, 364
aims 328
analytical chemistry 330
definition 328
development tools 329
structure 329, 330
explained variance principal components 156

extended *t*-test and *F*-test 37
externally studentized residual 270

f

factor
  coding 132, 133
  experimental 15, 143
  loadings 164, 165
  rotation methods 161
  score determination 169
  score matrix 169
  significance 102
factor analysis (FA)
  evolving 177
  factorial methods 177
  parallel 179, 279
  vs. singular value decomposition 170
factor effects 52, 101, 114, 116, 120–122, 129, 130, 133, 135
  estimation 120–122, 124
  vs. regression parameters 133
  relation with regression parameters 252, 268
factorial
  experiments 113, 116
  methods 151, 155, 182, 336
    factor analysis 161, 177
    principal component analysis 152
    singular value decomposition 161
factorial methods 86
fast Hadamard transformation (FHT) 79
feature 33, 90, 146, 148, 152, 159, 163, 164, 168, 171, 172, 186–189, 194, 196, 198, 199, 203, 213, 215, 217, 218, 274, 298, 310, 312, 319, 345, 366, 378, 379
figure-of-merit. *See* objective function
filter
  high-pass 80
  Kalman 61, 68, 69
  low-pass 80
  moving average 62, 63
  polynomial 61, 68, 71, 72, 86
  recursive 68
  Savitzky–Golay 64, 67, 71, 72
  width 62–64, 67, 100
filtering algorithm 69
fixed-size simplex optimization method 137, 138
flowing system 10
forecasting. *See* prediction
FORTRAN 8
forward directed learning algorithm 342
Fourier transformation (FT)
  conventional spectroscopy 77
  convolution and deconvolution 80–82
  data reduction and background correction 82
  discrete 77
  integration 82
  inverse 77
  Nyquist frequency 76
  signal filtering 80
  time domain 76
fractal structures 13
fractional factorial designs 115, 118
fragmentation codes 301–303, 314
  HORD code 302
  HOSE code 301, 302
  structure generator 303
frame 324, 379
frequency
  aliased 77
  function 18
F-tests 36
full factorial design 114, 115
fuzzy arithmetic 364
fuzzy clustering 194
fuzzy inferences 364
fuzzy inference systems 360
fuzzy logic 360, 364
fuzzy set 194, 314, 321, 352–354, 358, 360, 364
  operation 358
fuzzy theory 12, 352
  applications 364, 365
  cardinality 359
  complement 358, 359
  intersection 358
  relative cardinality 359, 360
  set theory 353
  union 358, 359
F value 36, 37, 47, 46, 50, 240, 241, 244

g

gas chromatography 145, 279, 300, 364
GC/MS 8
Gaussian distribution 17, 18. *See also* normal distribution
Gaussian elimination 251
Gauss–Jordan elimination 251
general least squares (GLM) 53
generic structure. *See* Markush structure
genetic algorithms 13, 321, 365
applications 369, 370
computer heredity 366
crossover 367
initial population 366
mutations 367
selection/reproduction 366
simulated annealing (SA) 370–372
tabu search (TS) 372, 373, 375
geometric mean 22
global search strategies. *See* genetic algorithms
glucose 16, 91, 97
good laboratory practice (GLP) 385
goodness-of-fit test 39, 239, 240, 272
χ^2 goodness-of-fit test 38, 39
graph theory 303
Grubb's test 44

h

Haar 86, 87
Hadamard transformation (HT)
data reduction and background correction 83
FHT 79
Walsh function 78
wavelet transformations 77
hair 146, 147, 156–159, 172, 193, 197, 199, 203, 204
Hamming distance 312
hard disk 8
hardware 6, 62
harmonic mean 22
hat matrix 268
Hebb learning. *See* learning, associative
Hebb's law 337
heredity 366
Hessian matrix 284
heteroscedasticity 245, 246
hexadecimal number 5
hidden layer 216, 331, 338, 340–343
hierarchical cluster analysis
centroid linkage 192
complete linkage 192
first reduced matrix 189
fourth reduced matrix 190
median linkage 193
second reduced matrix 190
single linkage 192
third reduced matrix 190
unweighted average linkage 192
Ward's method 193
weighted average linkage 192
hierarchically ordered ring description (HORD) code 302. *See also* HORD code
hierarchically ordered spherical description of environment (HOSE) code 301, 302. *See also* HOSE code
hierarchical search trees 311
histogram 17, 18, 245
hit list 311
homoscedasticity 246
horn clause 328
householder reduction 251
human being 9, 322
3-hydroxypropanal 305
hypotheses testing 20
comparison of mean with true value 31
comparison of two means 35, 36
comparison of variances 36
error of first and second kind 41
distribution. *See* hypotheses distribution
mean *vs.* true value 35
steps 30
hypothesis
alternative 30–32, 34, 40
null 40, 41, 43, 44, 46, 52, 251, 252, 270
testing 28, 30, 34, 35, 41
hypothetical distribution
χ^2 goodness-of-fit test 39
Kolmogorov–Smirnov's test 39

i

ICA. *See* Independent Component Analysis (ICA)
identification 308, 343, 365
identity matrix 69, 165, 207
image processing 11
implication 325, 330
in-breadth searching 326
Independent Component Analysis (ICA) 177, 178
in-depth search 330
indicator function 171
induction 326
inferences 11, 15, 16, 325, 360, 364
  abduction 325, 326
  deduction 325
  engine 330
  induction 326
  logical connectives 325
  rule 325
influential observation 246, 270, 271, 275
information and system theory 11
infra red spectroscopy 231, 259
  NIR 231, 257, 259, 272, 273, 343
initial population 366, 368
integration algorithm 82
interface
  lEEE-488 7
interference noise 80
interlaboratory comparison 36, 47, 49
interpolating splines 84, 85
interpretable factors 161
interquartile range 22, 23, 25
intersection 309, 314, 358, 377, 359
inverse sensitivity 104
inverted lists 309–310
  library search 309, 311
I/O system 8

j

jackknifed residual 270, 274, 275. *See also* studentized residual
Jacobian matrix 284
JCAMP/DX exchange file 297–298
joint committee on atomic and molecular data (JCAMP/DX) 297

k

Kalman filter 68, 69
Karhunen–Loéve expansion 159
kernel functions 215
kinetic–enzymatic oxidation 120, 123
K-matrix approach 264–266
k-nearest neighbor (*k*-NN) method 192
  classification 208, 210
knowledge
  processing 12
  acquisition 12, 329
  base 326, 329
  processing 321
  representation 323, 324
knowledge processing and soft computing. *See* artificial intelligence (AI); expert systems; fuzzy theory; genetic algorithms; neural networks
known factors 113
Kohonen networks 336, 345
Kolmogoroff–Smirnov's test 39
kurtosis. *See* distribution

l

laboratory-information-and-management system (LIMS) 8, 308–309
lack-of-fit test 240
Lagrangian theory 214
lasso regression 276
latent variable 153, 227, 254, 288
Latin square. *See* experimental design
Latin square factorial design 118, 119
learning
  algorithm 332, 338, 341, 342, 346, 370
  associative 337, 339, 347
  competitive 337, 338, 346
  law 337, 338, 377
  paradigm 336
  supervised 336, 339, 343, 364
  unsupervised 145, 151, 185, 186, 336, 345, 364
Leary criterion 108
legal inference 325, 326
level attained. *See p*-level
library file 297
library search 207, 295, 309–316, 365

hierarchical search trees 311
inverted lists 309–310
sequential search 309
similarity measures for chemical structures 314–316. *See also* similarity measures for spectra
library spectrum 311–313
limit of determination 105, 380, 386
limited selective method 107
linear discriminant analysis (LDA) 201, 225, 226
linearization method 282–285
linear learning machine (LLM) 198, 200
linear model 232, 240, 283
linguistic modifiers 356, 357
linguistic variables 356, 357, 361
linkage
centroid 192
complete 192
median 193
single 192
unweighted average 192
weighted average 192
liquid chromatography 128, 173
HPLC 111, 129, 139, 173, 281, 289, 290, 299, 312, 330, 343, 370
liquid handling 10
list processing language (LISP) 8, 322, 326–327, 330
local area networks (LANs) 8
location parameter 21
logical connective 5, 312, 325

m

Machine code 8
machine learning 12, 322, 323, 330, 347
Mahalanobis distance 186, 207, 269
Manhattan distance 186, 312
manipulation 9, 12, 61, 309
software 308
MANOVA. *See* analysis of variance, multidimensional
Markush chemical structures 307
Markush structure 307
Marquardt algorithm 285
mass spectrometry 88, 89, 145, 182, 279, 300, 312, 313, 323
electrospray–ionization (ESI-MS) 182, 184, 185
matrix inversion 266
mean
arithmetic 18, 21, 24
centering 148
geometric 22, 24
harmonic 22
c-means algorithm 364
measurement 2, 4, 8, 10, 11, 15, 23–25, 27, 28, 30, 31, 40, 48, 49, 51, 103–106, 109, 111, 120, 130, 232–236, 240, 247, 253, 279, 295, 300, 379, 382
model 68, 77, 83, 84, 91, 93, 95, 98, 99
uncertainty 16
median 23, 24, 193, 382
membership functions 194, 353–356, 358, 360, 362
characteristic function 353
discrete 357, 358
fuzzy observations 354
linguistic variables 356
truth values 357
memory
random access (RAM) 7, 310
read-only (ROM) 7
method of steepest descent 282, 285
Mexican hat 86, 87, 346, 347
Minkowski distance 187
missing data 147
mixture design 128
mode 3, 23, 163, 183, 279
model adequacy test 116, 233, 234, 236, 239, 240
model(ing)
empirical 124, 130, 131, 231, 281, 284, 290
linear 116, 130, 232, 240, 283, 336
mechanistic 130, 284, 290
power 211, 212
moments of distribution 19
Monte Carlo method 285
Moore–Penrose-matrix. *See* pseudo inverse matrix
Morgan's algorithm 307
Morlet 86, 87
multicomponent analysis 69, 107, 108, 249, 259, 267, 366, 370, 372

multicriteria optimization 364, 366
multidimensional analysis of variance (MANOVA) 51
multilayer perceptrons 341, 342
multiple linear regression
  diagnostic statistics. *See* diagnostic statistics
  direct calibration method 261, 263
  indirect calibration method 263, 264
  OLS. *See* ordinary least squares (OLS) regression
  PCR 249, 251, 253, 266, 274
  PLS 249, 251, 253, 255, 257, 266, 267, 274, 276
  P-matrix approach 265
  soft modeling 265, 266, 268
  spectrometric multicomponent analysis 249
multivariate adaptive regression splines (MARS) 281, 288, 289
multivariate curve resolution 155
multivariate regression spline model 288
multi-way analysis of variance 51
multiway decompositions 179
mutation 366, 367

n

negation 309, 325
Nelder and Mead method 101, 136
network 3, 8, 13, 215, 281, 295, 300, 308, 312, 333, 334, 336–339
  worldwide 8
Network Common Data Format (Netcdf) 300
neural networks 13, 312, 330, 370
  aggregation 332
  applications 343
  architecture 338, 339
  artificial neuron 331
  associative learning law 337
  backpropagation algorithm 341
  bidirectional associative memory (BAM) 333
  competitive learning law 337, 338
  convolutional 349
  deep learning 348
  hidden layer 331
  Kohonen network 345
  learning paradigm 336, 337
  long short-term memory (LSTM) 349
  models 336
  perceptron models 339
  recurrent 348
  transfer functions 332, 333
  weights of synapses 332
neuron 331, 336, 341–343, 345–347
nitrate 33–35
noise
  interference 80
  signal derivatives 71, 72, 74
non-hierarchical cluster analysis
  c-means clustering 194–196
  fuzzy clustering 194
nonlinear iterative partial least squares (NIPALS) algorithm 153, 154, 160, 210, 211, 255
nonlinear partial least squares (NPLS) 281, 287–290
nonlinear regression (NLR)
  chemical equilibria 281
  grid search method 285
  Hessian matrix 284
  Jacobian matrix 284
  Marquardt algorithm 285
  Monte Carlo method 285
  parameters 281
  regression diagnostics 285
  simplex method 285
  steepest descent 282, 283
nonparametric test 231, 281, 289, 290
normal distribution 18, 20, 21, 24, 28, 39, 40, 43, 205, 206, 210, 245
normalization 27, 64, 149, 227, 283
nuclear magnetic resonance (NMR) 77, 297, 299, 316, 317, 323
null hypothesis 31–36. *See also* hypothesis
N-way PLS (N-PLS) 280
Nyquist frequency 76

o

object 22, 147, 152, 157, 160, 186, 188–191, 194, 195, 203, 205, 207, 208, 210–213, 384
objective criterion 11, 109, 140, 382, 383
  aggregation of 189

objective function 103, 282, 285, 337, 364, 366, 367, 369, 372
accuracy of analysis 106
aggregation, performance characteristics 109, 110
calibration function 104, 105
detection limit and limit of determination 105
polyoptimization 110, 111
specificity and selectivity 106–108
time, cost and risk 108
oblimax criterion 172
one-variable *t*-test
calculated value *vs*. tabulated value 32, 33
nonparametric tests 33
null hypothesis 31
one-sided *t*-test 33
p-level 35
significance level 31
standard deviation 31
two-sided *t*-test 33
one-way analysis of variance 45
OPLS. *See* orthogonal PLS (OPLS)
optimization
poly 110
sequential 103, 135
simultaneous 109
systematic 101, 102
ordinal scaled values 104
ordinary least squares (OLS) regression
definition 250
matrix notation 250
parameter estimation 250
parameter significance 251, 252
prediction 252, 253
singular value decomposition (SVD) 251
Organization for Economic Cooperation and Development (OECD) 385
orthogonal PLS (OPLS) 257
outlier 15, 23, 25, 43, 44, 192, 200, 205, 211, 212, 246, 249, 267, 269–271, 275
Dixon's Q-test 43
Grubb's test 44
out-of-bag classification error 224
overfitting 179, 218, 251

p

parallel factor analysis 179, 182
parameter estimation 125, 130, 131, 133, 146, 233, 241, 249, 250, 252, 284, 285, 337, 343, 366
Pareto optimality 110
Pareto scaling 148
partial least squares (PLS) 153, 198, 249, 251, 257, 267, 276
regression 253–256
partition 17, 45, 114, 194–196, 216, 217, 291
pattern 194–196, 216, 217
pattern recognition 11, 12, 43, 145
pattern recognition and classification. *See* cluster analysis; supervised learning methods; unsupervised learning methods
PCR. *See* regression, principal component
peak separation after Kaiser 109, 110
peak shape
Gaussian 67
Lorentz 67, 72
peak list 299
Pearson distance 186
percentile 23
perceptron 216, 339, 343, 377
performance characteristic 104, 109, 110, 308, 380, *See also* objective function
periodicity 93, 95
pH monitoring 91
pharmaceutical industry 128, 259
phenol 30
phosgene 303, 304, 306
phosphate 188
photomultiplier 3, 20, 80
p-level 35, 50, 52, 272
Poisson distribution 20
polycyclic aromatic hydrocarbons (PAHs) 43, 173
polyoptimization 110, 111
potassium 46, 47, 51
Powell's method 139
p-phenylenediamine (PPD) 120, 123, 130–133, 140, 142
precision 10, 16, 101, 102, 106, 114, 251, 255, 256, 263, 380, 381
predicate logic 323

prediction 2, 69, 93, 99, 176, 218, 219, 233, 243, 244, 247, 252, 253, 255, 257, 267, 268, 272, 274, 287, 291, 345, 347, 370
  error 269, 274, 291
principal component 157, 210, 211, 228, 249, 255, 257, 263, 267–269, 276, 287, 312
principal component analysis (PCA) 150–154, 157, 158, 160, 162, 164, 165, 168, 169, 171, 177–179
principal component regression (PCR) 253, 266, 274
principal factor analysis 164, 168
probability 16, 18–20, 24, 27, 28, 30, 37, 206, 207, 240, 252, 353, 354, 367–370, 372
probability density function. *See* frequency function
probability function of Poisson distribution 20, 370
productivity 10
programming 3, 8, 10, 322, 326, 328
programming in logic (PROLOG) 322, 327–330, 377
  language 8, 322, 328–330, 377
protein 88, 90, 182, 185, 231, 257, 259, 271–274, 300
  apomyoglobin 183
proximity matrix 224
pruning 217
pseudo-inverse matrix 253

q

Q analysis 162, 163, 170
quadratic discriminant analysis (QDA) 207
quality assurance
  external 384
  internal 381–382
quality control 11, 365, 379–383
quality, definition 380
quantile 21, 23, 31, 32, 381
quartimax criterion 172

r

R analysis 162, 163, 166, 170
radial basis function 215
random access memory (RAM) 7. *See also* memory, random access
random errors 16, 106
randomization 112, 113, 143
random forest 222
random number distribution
  analytical 380, 381
  central limit theorem 21
  classes 16
  dynamic 380, 381
  frequency distribution 17
  Gaussian distribution 17, 18
  moments of distribution 19
  Poisson distribution 20
  scaling 148
  spectrophotometric measurements 17
random process 97
range
  analytical 80, 104
  dynamic 105
rank 21, 104, 165, 171, 269, 274
rank annihilation 151
read-only memory (ROM). *See* memory; read-only (ROM)
recovery function 273, 285, 292, 380
recursive filter 68, 69
reduced correlation matrix 163, 164, 168
regression
  linear 86, 97, 99, 173, 176, 202, 231, 232, 234, 235, 238, 245, 246, 284, 285, 316, 372, 380
  multiple linear 173, 176, 231, 249, 316
  multiway 278–280
  nonlinear 231, 281–283, 287, 289, 290
  nonparametric 290
  ordinary least squares 249–251, 253, 263, 266, 272–274
  partial least squares 253–255, 257
  principal component 249, 253, 266, 274
  robust 248, 249
  tree-based regression 291
  univariate linear 232
  weighted 245–248
regression diagnostic 231, 246, 249, 267, 273, 285
regression spline. *See* spline
regression trees 290, 291
relative cardinality 360
replication 111, 114, 125, 126, 128, 236

residual
analysis 245, 246
studentized 270, 271
resolution
analytical 107
chromatographic 109
visual 62, 71, 72
response 51, 52, 101, 102, 109, 111, 116, 124, 126, 128–130, 132, 135, 137, 139, 141, 142, 200, 216–218, 231, 252, 280, 281, 291, 367, 368
response surface designs
factor effects *vs.* regression parameters 133, 134
blocking of experiments 134
Box–Behnken design 126–128
central composite design 125, 126
mixture design 128, 129
response surface methods 129–132
response surface methods (RSM) 101, 103, 129, 130, 132, 231
ridge regression 276
risk 24, 29–31, 36, 41, 44, 52, 109, 205, 232, 240, 252, 270
robot
geometry 9
robotics and automation 9
robustness 101, 249, 270, 380, 381, 386
robust regression 248
round-robin test 381, 384
rotation
oblique 171
orthogonal 172
target 172
rotation of factors 171, 172
round-robin test 49
rule 2, 3, 10, 15, 30, 64, 74, 75, 78, 123, 153, 183, 205, 213, 299, 306, 324, 325, 328, 330, 337, 360
rule-based fuzzy systems 360–362

S

sample 6, 15, 91, 107, 108, 145, 233, 296–298, 357, 366, 380
Savitzky–Golay filter
filter coefficients 64
signal derivatives 71, 73, 74
scaling 86, 88, 104, 114, 148, 154, 172, 186, 249, 270, 379
auto 148, 149, 227
methods 148
range 148, 227
score 152, 154, 157, 158, 160–162, 168, 169, 174, 203, 211, 227, 254–257, 268, 269, 287
Scree plot 157
scree test 156, 157, 171
scree test principal components 157
screening designs
factor effects estimation 114, 120
fractional factorial designs 115, 120
full factorial design 114
search methods 103, 282, 285, 295
search problems. *See* genetic algorithms
search strategies 295, 309, 326
seed 186
selection 64, 67, 102, 113, 246, 309, 312, 323, 330, 339, 343, 366, 368, 374, 378
selective analytical method 106
selectivity 102, 106–110, 139, 259, 384
self-organizing feature maps 345. *See also* Kohonen networks
self-organizing nets. *See* Kohonen networks
semantic nets 324
sensitivity 101, 102, 104, 113
separation 2, 10, 109, 110, 139, 178, 201, 203, 207, 208, 213, 215, 281, 282, 300, 330, 339, 340
sequential optimization 103, 136
sequential search strategy 309
sequential strategies 103
shell 8, 329, 330, 377
shot noise. *See* noise, white
signals 74
derivative 71, 73, 74
Gaussian 178
integration 74, 75
Lorentz 84
processing 11, 61, 99
signal smoothing 84
signal-to-background ratio 108, 109
signal-to-noise ratio 80, 101, 108

significance
  level 30, 33, 35, 36, 39, 40, 43, 45, 50, 240–242, 244, 252, 272, 381, 382
  of parameters 251, 252
significant factors 171
similarity measure 186, 187, 311
similarity measures for spectra
  correlation measures 311
  dissimilarity measure 312
  grouping and feature selection 312
  logical operations 312
simplex method 11, 101, 136, 137, 139, 285, 371, 377
  fixed-size 137–138
  optimization method 138
  variable-size 135–141
Simpson rules 74, 75
simulated annealing (SA) 370
simultaneous strategies 102
singular value decomposition (SVD) 151–153, 160, 170, 251, 253, 265, 266, 280
skewness 19, 21, 24
smoothing algorithm 287
smoothing splines 86
soft computing 12, 321
soft independent modeling of class analogies (SIMCA) 198
soft modeling 265–267
software 2, 6, 10, 35, 40, 51, 52, 62, 118, 142, 157, 168, 267, 308
soil 43, 91, 228
source file 297
specific analytical method 107
specificity 107
spectra interpretation 12, 330
spectra simulation 316–318
spectrometric multicomponent analysis 259
spline
  adaptive 85
  function 288
  interpolating 84
  regression 281, 288–290
  smoothing 61
spline functions 84
  defined 85
  interpolating splines 85
  smoothing 86
standard deviation
  of the procedure 380
  relative 21, 23, 24
standard error of calibration (SEC) 268, 272
standard error of prediction (SEP) 256, 269, 273
standard normal distribution 28, 31. *See also* normal distribution
standard normal variate (SNV) 259
standard operating procedure (SOP) 381
standard solution 294, 380, 381
Staphylococcus nuclease (SNase) 89
star and sun-ray plots 196, 197
star design. *See* experimental design
star plot 197
stationary process 94, 97
statistical testing
  hypotheses. *See* hypotheses testing
  normal distribution 31
  outlier. *See* outlier testing
statistical tests 28, 33, 35, 43, 101, 102, 113, 234
statistics
  descriptive 16, 96
  inference 11, 15, 16
steel 37, 49, 51, 309, 323, 330
storage
  bulk 7, 297, 298
  optical 3, 20, 52, 179, 339
straight-line model. *See also* regression, univariate linear
  ANOVA 234, 235, 236, 238
  coefficient of determination 239
  confidence bands 243, 244
  confidence intervals 241, 242, 252
  correlation coefficient 239
  lack-of-fit test 240
  parameter estimation 233, 250, 251
  prediction 233, 252
  residual analysis 245
  residual variance 233
  robust regression 248
  standard deviation 233, 247, 283
  standard generalization 233
  weighted regression 246–248

structure–activity relationships 2, 318, 375
structure generator 303
Student's *t* value 31, 34, 121, 252, 270
studentized residual 270, 271
substructure search 307, 314
sulfate 93, 217
sum of residuals 238
sum of squares
  mean (MSS) 234, 237, 239
  predictive residual (PRESS) 269, 272, 274, 276, 289
sun-ray plot 197
supervised competitive learning 338
supervised learning methods 336, 339, 343, 364, 365
  discriminant analysis 201–207
  k-nearest neighbor method 207, 208
  linear learning machine 198–200
  soft independent modeling of class analogies 210, 211
  support vector machines 215, 216
  tree-based classification 216
supervision 10
support vector machines (SVMs) 198, 213, 215, 216
symbolic knowledge processing 360, 377
  inferencing 325
  knowledge representation 323, 324
  programming 326, 328
  search strategies 326
symbolic programming 322, 323, 326
symmetry 125
Symmlet 87
systematic errors 16, 105
systematic optimizations 101, 102

t

tabu search (TS) 372
tail area 28, 30
target-transform factor analysis (TTFA) 162
Taylor expansion 284
ternary mixtures 129
test
  chi-square 38
  distribution 15, 20, 30, 31, 38
  F 33
  for distributions 38
  Kolmogorov–Smirnov 38, 40
  nonparametric 33, 207
  outlier 43
  t-, extended 37
  t-, general 30, 36
  t-, one-sided 32–35
  t-, one-variable 31
  t-, two-sided 15, 33, 35, 37
  t-, two-variable 35
  Wilk's 52
textile industry 128
time-series analysis 16, 61, 99
  autocorrelation. *See* autocorrelation
  autocovariance 94
  autoregression 99
  cross-correlation 98
titanium 36, 37
transfer function 332, 333, 377
transformation
  Fourier 11, 76, 227, 297
    discrete 77
    inverse 77
  Hadamard 11, 61, 76–79
  wavelet discrete (DWT) 88, 89
traceability 384
trapezoidal rule 74
tree-based classification 217
  CART 216–220, 225, 229
  classification 216
  ensemble methods 218, 219
  recursive binary partitioning 216
trueness 10, 16, 101, 106, 380, 381, 384
truth value 5, 16, 31, 33, 106, 254, 328, 357, 358, 384
Tucker models 179, 183, 185, 279
two-variable *t*-tests 35
two-way analysis of variance 48, 50

u

uncertainty 15, 16, 25, 27, 28, 69, 354, 355. *See also* measurement uncertainty
  vs. error 28
  error propagation 25
  measurement uncertainty 16
  random and systematic errors 16
  sources 26

uncertainty (*contd.*)
vagueness 16
variability 15
uncontrolled factors 113
undirected graphs 303
univariate linear regression 232
straight-line model 232, 233. *See also* straight-line model
unknown factors 113
unselective or partially selective method 107
unsupervised competitive learning 338
unsupervised learning methods 336, 365
cluster analysis 185
factorial methods 151, 152
graphical methods 196
urine 97
UV spectroscopy 173

V

vagueness 15, 16
validation 10, 86, 156, 157, 362, 380, 384
variability 15, 16, 364
variable-size simplex optimization method 138
variance
explained 156
meaningful 212
residual 157, 211–213, 233, 241
variance–covariance matrix 150. *See also* covariance matrix
variate 20, 28
standard normal 28, 31
varimax criterion 172

W

Walsh function 77, 84
Walter Shewhart 383
Ward's method 193
warping 91
water 30, 33, 35, 46, 51, 91, 129, 231, 259, 311, 356
wavelet analysis 86, 88, 90
weighing 10, 279, 280
weight 43, 74, 109, 128, 199, 200, 214, 219, 246, 247, 255, 257, 331–333, 335–338, 347, 361, 377
weighted regression 245–248
Welch test 36
Wilcoxon test 33
Word 3, 6, 32, 152, 240
working range 104, 105. *See also* range, analytical

X

X-ray analysis 20
X-ray diffraction spectrometry 297
X-ray fluorescence analysis (XRF) 109, 323, 353, 354, 365, 370
XOR problem 312, 340