Index

μ-X-ray fluorescence, μ-XRF, see glass, μ-X-ray fluorescence

accuracy, 192, 195, 204
analysis of variance, see ANOVA
annealing, see glass, annealing
ANOVA, 78–85
arithmetic mean, see data, central tendency
artificial neural network, ANN, see classification problem, artificial
neural network
Bayes’ theorem, 31–32
between-object variability, 26, 42–44, 107–180
box-plots, 44, 114, 117, 121, 126, 131, 138, 159, 162, 165, 170, 258
calibration, 192, 195
car paints, 10–13
    likelihood ratio, LR, 119–125, 205–209
Py-GC/MS, 119–125
central tendency, see data, central tendency
classification problem, 2, 8, 20, 27–31, 32, 82–84
    artificial neural network, ANN, 28–30
discriminant analysis, DA, 27, 29
likelihood ratio, LR, 31, 151–180

naïve Bayesian classifier, NBC, 28, 31
partial least-squares discriminant analysis, PLS-DA, 28
SIMCA, 28, 29, 88
support vector machines, SVM, 28, 31
cluster analysis, CA, 28, 85–92
    averaged linkage method, 89
    clustering, 89–91
dendrogram, 89–91
    furthest neighbour, 89
    hierarchical cluster analysis, 89–92
    nearest neighbour, 89
    Ward’s algorithm, 89–91
comparison problem, 2, 12, 13, 20, 21–27, 32
    likelihood ratio, LR, 23–26, 107–150
two-stage approach, 21–23, 26–27, 69–71
confidence interval, 67–69
data
    central tendency, 39
        arithmetic mean, 39–41, 257
        median, 39–41, 257
        mode, 39–41
        continuous, 35–105
covariance, 45–49
descriptive statistics, 39–59
discrete, 23, 35, 107, 119, 151, 225–227, 296–300, 305
INDEX

data (continued)
dispersion, 42–44
  interquartile range, IQR, 45, 257
  standard deviation, 42, 257
  variance, 42, 49, 257
location, 39–41
  interquartile range, IQR, 45, 257
  quantiles, 40, 44, 257
  quartiles, 41, 44
normal distribution, see normal distribution
Pearson’s correlation coefficient, 45–49, 52, 93, 263–265
correlation significance test, see hypothesis testing, correlation significance test
transformation, 25, 30, 37–38
  autoscaling, 37, 46, 47, 90, 93, 94
  hyperspherical, 38
  logarithmic, 37, 120, 130, 137, 165
variance–covariance matrix, 48, 99, 103, 132, 134
degrees of freedom, 42, 43, 53–56, 63, 64, 69–71, 73, 74, 77, 80–82, 261
descriptive statistics, see data, descriptive statistics
DET plot, 190
dimensionality reduction, 92–105, 131–137
graphical models, 99–105, 131–137
principal component analysis, PCA, 93–99
discriminant analysis, DA, see classification problem, discriminant analysis
discriminating power, 185
dispersion, see data, dispersion
distance
  between objects, 85–88
  between variables, 86, 88
  Chebyshev distance, 87
distance matrix, 91
  Euclidean distance, 30, 87, 90
  Mahalanobis distance, 29
  Manhattan distance, 30, 87
  square Euclidean distance, 87
distribution
  F distribution, 55
t distribution, 53–54
chi-squared distribution, 54
standard normal distribution, 49–53, 260
three-sigma rule, 51
drug profiling, 30–31
t-distribution, 52–54
empirical cross-entropy (ECE), 192, 197
t-distribution, 53–54
t-distribution, 52–54
empirical performance, 182
estimators, 39
evaluation of the evidence, 2, 19–32
fibres, 13–15
  likelihood ratio, LR, 116–119
  MSP-DAD, 116–119
files
calculatoR.calculatot.exe, 310
codes
car_paints_histogram_Tippett_DET _ECE_code.R, 205
comparison_research_code.R, 144–149, 274–278
dR classification_code.R, 163
GCMS_car_paints_comparison _code.R, 121, 123
MSP_inks_comparison_code.R, 113, 115, 128
NaSiCa_model_SEMEDX_histogram _Tippett_DET_ECE_code.R, 210
SEMEDX_glass_classification _casework_code.R, 171
SEMEDX_glass_classification _code.R, 166
SEMEDX_glass_comparison _casework_code.R, 139
SEMEDX_glass_comparison_code.R, 136
x_y_z_MSP_inks_histogram_Tippett _DET_ECE_code.R, 201
databases
ANOV A_data.txt, 81, 82
cwp_database.txt, 172, 173, 175–177, 282–285
data_control.txt, 140, 141, 271
data_recovered.txt, 140, 271
database_CA.txt, 90–92
database_PCA.txt, 94–96
dRI_c_p_database.txt, 163
dRI_c_w_database.txt, 163
dRI_cwp_database.txt, 161, 162
dRI_w_p_database.txt, 163
evidence.txt, 172, 173, 282–285
GCMS_car_paints_database.txt, 120–123
GCMS_flammable_liquids_database .txt, 158, 159
glass_data.txt, 36, 48, 49, 97, 98, 108, 253
glass_data_package.txt, 45
glass_database.txt, 141, 144, 145, 273–278
MSP_fibres_database.txt, 117
MSP_inks_database.txt, 113, 114, 127
RI_1.txt, 71–72, 85
RI_2.txt, 71–72, 85
RI_annealing.txt, 74
RI_control.txt, 125–127, 311, 312
RI_database.txt, 26
RI_population_Wiley.txt, 311–313
RI_recovered.txt, 125–127, 311, 312
RI_w_database.txt, 125, 126
SEMEDX_control.txt, 137, 138
SEMEDX CW_p_database.txt, 165–169
SEMEDX_evidence.txt, 169, 170
SEMEDX_glass_database.txt, 99, 130–133
SEMEDX_recovered.txt, 137, 138
T2_Hotelling.txt, 76
variables_selection.txt, 83, 84
functions
DET_function.R, 201, 205, 210, 291–292
ECE_function.R, 201, 205, 210, 292–293
histogram_function.R, 201, 205, 210, 289–290
LR_classification_one_level_KDE.R, 163, 167, 171, 173, 176, 282–289
LR_classification_one_level_Nor.R, 173, 176, 282–289
LR_classification_two_level_KDE.R, 163, 167, 171, 173, 176, 281–289
LR_classification_two_level_Nor.R, 173, 176, 280–289
LR_comparison_KDE.R, 122, 136, 139, 142, 144, 149, 269, 271, 273
LR_comparison_Nor.R, 115, 128, 142, 144, 149, 269, 271, 273
scaled_inverse_cov_matrix.R, 99
Tippett_function.R, 201, 205, 210, 290–291
UC1_one_level_calculations.R, 163, 166, 171, 173, 176, 282, 285
UC1_two_level_calculations.R, 163, 166, 171, 173, 176, 278–280, 282, 285
UC2_one_level_calculations.R, 163, 166, 171, 173, 176, 282, 285
UC2_two_level_calculations.R, 163, 166, 171, 173, 176, 282, 285
UC_comparision_calculations.R, 44, 113, 121, 128, 136, 139, 141, 144, 149, 266–268, 273–278
graphics
BMA_research_descriptive_statistics .eps, 124
BMA_research_LR_distribution.eps, 124
dRI_c_p_research_descriptive .statistics.eps, 164
dRI_c_w_research_descriptive .LR_distribution .eps, 164
dRI_c_w_research_LR_distribution .eps, 164
files (continued)

dRI_w_p_research_descriptive
_statistics.eps, 164
dRI_w_p_research_LR_distribution
.ep, 164
I16_research_descriptive
_statistics.ep, 124
I16_research_LR_distribution.ep, 124
logAlO_cw_p_classification
_casework_Nor.txt, 174
logAlO_cw_p_classification
_casework_KDE.ep, 174
logAlO_cw_p_classification
_research_Nor.ep, 177
logAlO_cw_p_classification
_research_KDE.ep, 177
logAlO_cw_p_casework_descriptive
_statistics.ep, 171, 174, 175
logAlO_cw_p_research_descriptive
_statistics.ep, 178
logAlO_cw_p_research_LR
_distribution_KDE.ep, 179
logAlO_cw_p_research_LR
_distribution_Nor.ep, 178
logAlO_p_cw_research_descriptive
_statistics.ep, 168
logAlO_research_descriptive
_statistics.ep, 137
logCaO_cw_p_casework_descriptive
_statistics.ep, 171
logCaO_cw_p_research_descriptive
_statistics.ep, 168
logCaO_research_descriptive
_statistics.ep, 137
logFeO_cw_p_casework_descriptive
_statistics.ep, 171
logFeO_cw_p_research_descriptive
_statistics.ep, 168
logFeO_research_descriptive
_statistics.ep, 137
logKO_cw_p_casework_descriptive
_statistics.ep, 171
logKO_cw_p_research_descriptive
_statistics.ep, 168
logKO_research_descriptive
_statistics.ep, 137
logMgO_cw_p_casework_descriptive
_statistics.ep, 171
logMgO_p_cw_research_descriptive
_statistics.ep, 168
logMgO_research_descriptive
_statistics.ep, 137
logNaO_casework.ep, 143
logNaO_cw_p_casework_descriptive
_statistics.ep, 171
logNaO_cw_p_research_descriptive
_statistics.ep, 168
logNaO_research_descriptive
_statistics.ep, 137, 147, 149
logNaO_research_LR_distribution
.ep, 148, 149
logSiO_cw_p_casework_descriptive
_statistics.ep, 171
logSiO_cw_p_research_descriptive
_statistics.ep, 168
logSiO_research_descriptive
_statistics.ep, 137
M2E_research_descriptive_statistics
.ep, 124
M2E_research_LR_distribution.ep, 124
M2P_research_descriptive_statistics
.ep, 124
M2P_research_LR_distribution.ep, 124
MMA_research_descriptive_statistics
.ep, 124
MMA_research_LR_distribution.ep, 124
model_DET.ep, 206, 210
model_ECE.ep, 206, 210
model_histogram.ep, 206, 210
model_p_cw_research_LR
_distribution.ep, 168
model_research_LR_distribution.ep, 124, 137
model_Tippett.ep, 206, 210
MST_research_descriptive_statistics
.ep, 124
MST_research_LR_distribution.ep, 124
NaSiCa_DET.ep, 210
NaSiCa_histogram.ep, 210
files (continued)
M2P_comparison_research_KDE_different.txt, 124
M2P_comparison_research_KDE_same.txt, 124
MMA_comparison_research_KDE.txt, 123, 124
MMA_comparison_research_KDE_different.txt, 124
MMA_comparison_research_KDE_same.txt, 124
MMA_comparison_research_KDE_same.txt, 124
model_comparison_casework_KDE.txt, 139
model_comparison_research_KDE.txt, 136
model_comparison_research_KDE_different.txt, 136, 205
model_comparison_research_KDE_same.txt, 137, 205
model_cw_p_classification_casework_KDE.txt, 171
model_p_cw_classification_research_correct_rate.txt, 168
model_p_cw_classification_research_KDE.txt, 168, 210
MST_comparison_research_KDE.txt, 123, 124
MST_comparison_research_KDE_different.txt, 124
MST_comparison_research_KDE_same.txt, 124
NaSiCa_p_cw_classification_research_KDE.txt, 210
RI_comparison_casework_KDE.txt, 126
TOL_comparison_research_KDE.txt, 123, 124
TOL_comparison_research_KDE_different.txt, 124
TOL_comparison_research_KDE_same.txt, 124
x_comparison_research_Nor.txt, 115
x_comparison_research_Nor_different.txt, 115, 201
x_comparison_research_Nor_same.txt, 116, 201
x_y_comparison_research_Nor_different.txt, 129
x_y_comparison_research_Nor_same.txt, 129
x_z_comparison_research_Nor.txt, 115
x_z_comparison_research_Nor_different.txt, 115, 201
x_z_comparison_research_Nor_same.txt, 116, 201
y_comparison_research_Nor.txt, 115
y_comparison_research_Nor_different.txt, 115, 201
y_comparison_research_Nor_same.txt, 116, 201
y_z_comparison_research_Nor.txt, 128
y_z_comparison_research_Nor_different.txt, 129
y_z_comparison_research_Nor_same.txt, 129
z_comparison_research_Nor.txt, 115
z_comparison_research_Nor_different.txt, 115, 201
z_comparison_research_Nor_same.txt, 116, 201
flammable liquids, 8–10
adsorption thermo-desorption gas chromatography, ATD-GC/MS, 8–10, 158–161
diesel fuel, 10, 31, 158–161
kerosene, 10, 31, 158, 161
likelihood ratio, LR, 158–161
forensic expert, xii, 19, 31
Gaussian kernel function, 57
glass, 2–8
μ-X-ray fluorescence, μ-XRF, 4
annealing, 7, 74–75, 161–164
classification problem, 161–172
comparison problem, 125–126, 129–139
glass refractive index measurement, GRIM, 5–8
laser ablation inductively coupled plasma mass spectrometry, LA-ICP-MS, 4
refractive index, RI, 26–27, 65–67,
69–72, 74–75, 85, 125–126,
161–164
scanning electron microscopy coupled
with an energy dispersive
X-ray spectrometer,
SEM-EDX, 4–5, 129–139,
164–172
glass refractive index measurement, GRIM,
see glass, glass refractive index
measurement
graphical models, see dimensionality
reduction, graphical models
ground-truth, 182
histograms, 45, 262
hypothesis testing, 26, 59–78, 265
$F$-test, 85
t-test, 26, 63–75
alternative hypothesis, 59–78
correlation significance test, 77–78
Hotelling’s $T^2$ test, 75–77
null hypothesis, 59–78
one-sided, 60, 61, 64, 69
paired $t$-test, 72–75
two-sided, 60, 64, 67, 69–78
Welch test, 71, 72
inks, 13–15
likelihood ratio, LR, 112–116, 127–129,
200–205
interquartile range, IQR, see data,
dispersion, interquartile range
kernel density estimation, KDE, 24, 25,
56–59, 110–112, 123, 133,
155–157, 168, 169, 263, 294, 309
laser ablation inductively coupled plasma
mass spectrometry, LA-ICP-MS,
see glass, laser ablation
inductively coupled plasma mass
spectrometry
likelihood ratio, LR
classification problem, 31, 151–180
application of KDE for between-object
data distribution, 155–157
normal between-object data
distribution, 152–155
one-level, 154, 156
two-level, 152–154, 155–156
comparison problem, 23–26, 107–150
application of KDE for between-object
data distribution, 110–112
normal between-object data
distribution, 108–110
logarithmic strictly proper scoring rule, 194
lower quartile, see data, quartiles, quartiles
matrix, 36, 228
column vector, 36
distance matrix, 89, 91
identity matrix, 37, 233
row vector, 36
symmetric matrix, 36, 37, 48
variance–covariance matrix, see data,
variance–covariance matrix
median, see data, central tendency,
median
misleading evidence, 184
mode, see data, central tendency, mode
MSP-DAD
CIELab, 15
CIEXYZ, 15
CIExyz, 15
colour, 13–15
fibres, 13–15
inks, 13–15
na"ıve Bayesian classifier, NBC, see
classification problem, naïve
Bayesian classifier
neutral method, 199
neutral reference, 199
normal distribution, 49–53, 56, 108–110,
113, 117, 128, 152–155, 159,
259, 260
oracle probability distribution, 195
partial least-squares discriminant analysis,
PLS-DA, see classification
problem, partial least-squares
discriminant analysis
pool adjacent violators algorithm, 199, 240–243
principal component analysis, PCA, see dimensionality reduction, principal component analysis
probability density function, 23–25, 49–59, 107, 134, 151, 259, 260, 263
Q-Q plots, 38, 53, 114, 117, 122, 126, 132, 159, 162, 166, 167, 259
quantiles, see data, location, quantiles
quartiles, see data, quantiles, quartiles
refractive index, RI, see glass, refractive index
reliability, statistics, 196
scanning electron microscope coupled with an energy dispersive X-ray spectrometer, SEM-EDX, see glass, scanning electron microscope coupled with an
energy dispersive X-ray spectrometer
significance level, 22, 61, 62, 66
smoothing parameter, 57, 58, 110–112, 155–157
standard deviation, see data, dispersion, standard deviation
support vector machines, SVM, see classification problem, support vector machines
two-stage approach, see comparison problem, two-stage approach
upper quartile, see data, quantiles, quartiles
validation, 182
validation database, 182
variance, see data, dispersion, variance
within-object variability, 26, 42–44, 107–180