# INDEX

Adaptive estimates, 150  
Adaptive trimmed likelihood, 160  
  location, 160, 164, 169  
  regression, 163–169  
Adaptive trimmed likelihood algorithm (ATLA), 163–169, 172, 176, 181  
estimator, 165  
Amemiya, T., 89  
Asymptotic  
bias, 49, 107  
  breakdown point, 144  
  covariance matrix, 72  
efficiency, 17, 18, 89, 144  
  efficiency of the median, 18  
normality, 15, 60, 73, 77, 81, 82, 92–94  
  unbiasedness, 77  
  variance, 49, 60, 61, 77, 144, 177  
Asymptotically efficient estimator, 69  
Asymptotically efficient root, 47  
Asymptotically unique consistent sequence of roots, 14  
Asymptotically unique consistent root, 10–11, 17, 32–33, 37, 38, 94–95, 110  
Autocorrelated data, 28  
Auxiliary functional, 63  
Beaton, A.E. and Tukey, J.W., 8, 20  
Bernoulli random variable, 6  
Bias  
  adjustments, 102  
  corrected scale estimator, 105  
  expansion, 99  
Big Data, 38, 182  
Bootstrap, 105–107  
  bias corrected estimate, 106  
Borel sets, 2, 28  
Bouligand derivative, 73  
Breakdown point, 61, 154, 158, 160  
  asymptotic, 125  
  of the trimmed likelihood estimator, 158, 164  
  of the trimmed mean, 158  

---  

© 2018 John Wiley & Sons, Inc. Published 2018 by John Wiley & Sons, Inc.  
Companion website: www.wiley.com/go/clarke/robustnesstheoryandapplication
Calcutta Statistical Institute, 77
Central limit theorem, 1, 13, 14, 59
Characteristic function, 115–117
empirical, 116
Compact derivative, 52–53, 107, 153
Conditions A, 31–33, 35, 89
Conditions A’, 63, 65, 66
Conditions W, 43, 145
Consistency, 32, 37, 40, 43, 60, 73, 82
global, 35–37, 40
of the jackknife variance estimates, 102
local, 35, 37, 81
multiple roots, 82
Convergence
almost sure, 4, 5, 7
in distribution, 13, 19, 41
in probability, 3–5,
with probability one, 4
uniform, 5, 9, 17, 84, 108
uniform almost sure, 12
uniform over a class of functions, 16
Cox regression model, 184
Cramér Conditions, 73
Cramér–Rao inequality, 17
Cramér–Rao lower bound, 44, 76, 120
Cramér–Rao theory, 158
Cushny and Peebles data, 163
δ-method, 73
Distance
Cramér–von Mises, 114–115, 117, 120
Hellinger, 120, 123–124
integrated squared error, 116
Kolmogorov, 47, 121–2
Kolmogorov metric, 28–29, 54–55, 63,
69, 72, 76, 89, 102, 121, 145
Kolmogorov–Smirnov, 28, 121–2
Lévy, 28–29, 54–55, 122
metric, 28
Prohorov, 28–30, 32, 44, 54–55, 63
Distribution
beta, 142
Cauchy, 47, 51, 89, 98, 116, 189
chi-squared, 140
contaminated, 17
contaminated normal, 18
cumulative parametric, 30
exponential, 51, 182
gamma, 139
k-component exponential mixture, 137
k-component univariate normal
mixture, 125
Lévy, 116
log normal, 163, 182
mixture of normals, 97
multinomial distribution, 51
multivariate normal, 124, 176–177
negative exponential, 182
normal, 8, 9, 51, 116
Poisson, 51, 92
standard normal, 18
Student’s t, 50–51
symmetric, 19, 70
Distribution function, 4
empirical, 16–17, 27, 29–30, 37, 144
parametric, 30
Domain of attraction, 12–13
Dvoretzky–Kiefer–Wolfowitz
inequality, 46
Efficiency, 18, 27, 30, 92
multivariate parameters, 51
relative, 18–19
Efficient score function, 30, 37, 47, 83
Epsilon contaminated neighbourhood,
31, 60
Equicontinuity, 8, 37, 43, 64
Equicontinuous class of functions, 37, 69
Ergodic sequences, 8
Estimators
ATLA see Adaptive Trimmed
Likelihood Algorithm (ATLA)
β-trimmed mean, 182
Cramér–von Mises, 114, 117, 121
Cramér–von Mises type, 119–121
Fréchet differentiable, 44, 52–53
Geometric mean, 161
Hellinger Distance, 120
Hodges–Lehmann, 119, 144
integrated squared error, 116–117, 121
least median of squares (LMS),
154–155, 159, 166, 168
least squares estimator, 154
least trimmed squares (LTS), 154–156,
159, 164–166, 169, 172, 175
asymptotic variance, 155, 156
breakdown bound, 164
$L_2E$-minimum distance, 124
L-estimators, 74, 150–153, 173
$L_2$-minimum distance, 117, 119, 121,
124–126
for mixing proportions, 126–130
for switching regressions, 130
maximum likelihood estimator (MLE), 29, 59, 60, 69, 76, 77, 87, 100, 113, 114, 119, 123, 137, 139, 140, 142, 184

MCVM, 114, 141–142

M-estimators see M-estimator(s)

minimum covariance determinant (MCD), 155, 175–178

minimum distance, 29, 121

minimum Volume Ellipsoid estimator, 159

MLE see Maximum likelihood estimator(s)

MM-estimators, 169–172, 181
differentiability, 169

regression, 170, 182

uniform convergence, 169

weak continuity, 169

R-estimator of location, 144

S-estimators, 172

superefficient, 77

r-estimators, 172

trimmed likelihood estimate (TLE), 156, 160, 165

trimmed likelihood estimator (TLE), 152, 156, 158–159, 164, 170–171, 181, 182

of location, 156–158, 170, 171

regression, 170–171

of scale of a normal distribution, 156–157

trimmed mean, 147–151, 158, 173

optimal, 149

weakly continuous, 148

Winsorized mean, 148–150, 173

Expected value, 4

Fisher consistency, 30, 37, 42, 49–51, 73, 99, 153–156, 182

Fisher information, 17–18, 50, 120

Fixed point theorem, 75

For all sufficiently large n (f.a.s.l.n.), 7, 37

Forward Search, 176, 179

Fourier transform, 116

Fréchet derivative, 44, 46, 48–50, 52–53

strong, 101

Fréchet differentiability, 46–47, 49, 55, 59, 61–63, 67, 73, 89, 107, 118, 125, 128, 137, 142, 148, 151, 159, 181–182, 185

MLE, 47, 50

Fréchet expansion, 46, 103

Fréchet space, 36

locally uniform, 74, 75

stochastic, 73

strong, 107

function of total bounded variation, 69

Gamma function, 140

Gâteaux derivative, 52–53

Gaussianity, 28

Generalized Jacobian, 62–63

Glivenko–Cantelli theorem, 16

Global maximize theorem, 85

Goodness of fit, 82, 114

criteria, 122

statistic, 82

Anderson–Darling, 121

Cramer–von Mises, 121

Kolmogorov–Smirnov, 121

Lévy, 122

Gross errors, 28

Gross error sensitivity, 49, 53–54, 57, 61, 107

Grubb’s-type outlier rule, 153

Hadamard derivative, 52

Hampel three part redescender, 20, 79

Hermite, 62

Huber’s

least favourable distribution, 60, 69

minimax solution, 19, 60–61, 79, 148, 151

Proposal-2, 22–23, 68, 102, 104

regression, 70

skipped mean, 74, 153, 159, 181

Identifiability, 36, 86

Identifiable parametric models, 30

Implicit function theorem, 182

Independent identically distributed (i.i.d.), 3, 5–8, 26, 55

Influence function, 48, 55, 61, 174, 181

median, 56

M-estimator, 48

Initial estimates, 24

Inverse function theorem, 33, 62

Ion channel, 136

Jackknife, 100, 105, 107

bias-corrected estimator, 100

consistency, 107, 109

variance estimate, 100–101
Kernel density estimator, 120, 124
Kullback–Leibler divergence, 123

Laplace distribution, 171
Law of large numbers, 3, 4, 13, 130
   strong, 1, 5–7, 9, 16–17, 19, 89
   weak, 5
Least upper bound, 33
Leverage points, 70–71
Likelihood disparity, 123
Lipschitz, 62, 65
   constant, 65
$L_1$-loss function, 60
Local maximum, 38
Logarithms, 161
Loss function, 60
   least squares, 59

Markov chains
   finite state continuous time, 136
Maximum likelihood estimator(s), 25, 29, 30, 44, 54, 57, 59–60, 123, 137
   Cauchy
      location, 83, 86
      location and scale, 47
   gamma, 140, 142
   Laplace distribution, 60
singularity problems, 114
Maximum likelihood type equations, 29
Mean value theorem, 14, 21, 45, 62, 67
   for nonsmooth functions, 62
Median, 10, 13, 17–19, 26, 38, 53–54, 74–75, 148–149
   absolute deviation (MAD), 20
   absolute deviation normalized (MADN), 20, 26
   of the log normal distribution, 163
M-estimator(s), 1, 8, 13, 29, 59–60, 83, 120, 153
   B-optimal, 140–142
   first order bias for scale, 104
   generalized (GM), 70
   their introduction, 59
   for linear models, 71
   of location, 8, 10, 59–60, 144
   non-robustness, 38
   regression, 70, 119
   of scale, 21
   with strong expansions, 75

weak continuity, 33
Method of least squares, 60
Method of moments, 29, 84, 140, 142
M-functional
   qualitatively robust, 43
   weakly continuous, 43
Minimum distance estimation, 37, 98, 113
Minimum divergence method, 123
Mixture estimation, 115
   $k$ component exponential distributions, 136–137
   $k$ component univariate normal distributions, 125
Modified data on wood-specific gravity, 167
Moment generating function, 115–116
   empirical, 115
Monte Carlo simulation, 94
Multiple roots, 10, 21–22, 38, 79–98
   asymptotics, 80
Multivariate Central Limit Theorem, 46
Multivariate normal distribution, 82, 124, 176–178
Multivariate outliers, 175–179

Neighbourhood, 39
   Lévy, 66
   Prohorov, 42–43, 66
   Newcomb’s measurements of the passage time of light, 163
Newton algorithm, 108–109, 159
   iteration, 109, 159
   one step improvement, 109, 159
   Newton–Raphson algorithm, 18
   iteration, 11–13, 19, 26
   Neyman’s Chi-square, 122–123
   Nonsmooth analysis, 21, 62
   Normal density, 9
   Normed linear space, 52
   Null sequence, 34, 66

Observation space, 2, 4
Order statistics, 147

Parameter space, 27
Pearson’s Chi-square, 123
Poisson regression, 184
Probability density function, 30
Probability space, 2
Product measure, 3
Pseudo values, 100
p-values, 185

Qualitative robustness, 61
Quantile, 53
Quenouille, M.H., 99

R, 170
Rademacher’s theorem, 62
Random sample, 18
Redescending estimator
  Bachmaier’s choice, 23–24, 26
  location and scale, 23
Regression diagnostics, 169
Rejection point, 61, 80
Residual, 70
Robustness of efficiency, 148, 184
Robustness of validity, 148, 184
Robust testing, 184
Root $n$ shrinking neighbourhoods, 54
Rounding, 28

Sample average, 4, 13, 18, 148
Schwarz inequality, 50
Score function, 123
Scottish hill races data, 168
S-divergence family, 123
Selection functional, 17, 21–22, 35–36,
  38–40, 42, 80, 82, 84–92
  auxiliary, 32, 38, 63, 66
  Hampel, 38, 83
  mooted, 88, 98
Sensitivity curves, 140, 141
Separable metrizable space, 30, 63
Shorth, 156
  $(1-\alpha)$-shorth, 156, 164
Slutsky’s theorem, 15, 46
Small sample bias, 99, 114
Span, 156, 161, 164
Stable laws, 115, 116
Stack loss data, 168
Standard deviation, 4
Standard normal variable, 15
Statistical functional, 27
Stieltjes, 62
Stochastic expansion, 81
Strong expansions, 74, 75
Student $t$-confidence interval, 148
Support vector machine, 73
Supremum, 12
Survival analysis, 184
Switching regressions, 130–135

Taylor series expansion, 104
Telephone data, 167
Topology of weak convergence, 122
Trimmed likelihood principle, 152
Tukey bisquare, 8, 10, 12, 15–16, 18–20,
  26, 29, 32, 43, 159
  weak continuity, 47
Unbiased estimate, 19
Unbiased estimates of mixing proportions, 128
Unique global maximum, 39
Upper semi-continuous, 63, 66

Variance, 4, 102
Variance component models, 184

Weak continuity, 32–33, 41–43, 61–63, 66,
  89, 118, 125, 128, 137, 142,
  148, 182
  uniform, 43
Weak convergence, 41
Weakly continuous root, 32,
Western Australian goldfield data, 161
White, 84, 85