

Subject Index

a

- admissible 24, 88, 184
- algorithm
 - for Bernoulli sampling 51
 - of the cube method 130, 132
 - for Poisson sampling 93
 - for simple random sampling with replacement 56
 - for systematic sampling
 - with equal probabilities 57
 - with unequal probabilities 90
 - for unequal probability sampling designs 86
- alternative estimator 35
- approach
 - design-based 1, 13, 272, 273, 275, 277
 - model-assisted 1
 - model-based 1, 8, 11, 231, 263–280
 - reverse 152, 347
- approximately unbiased 203, 237
- approximation
 - bias 49
 - of the variance 111
- asymptotic statistic 295
- auxiliary
 - information 2, 9, 10, 32, 65, 72, 77, 83, 111, 120, 143, 195, 196, 200, 201, 209–211, 217, 225, 226, 237, 255–257
 - variable 83, 84, 100, 111, 123, 128, 131, 135, 141, 195, 199, 200, 202, 331
- b**
- balanced
 - sample 119–142
 - sampling 119–142
- balancing 119–142, 172
 - constraint 131
 - equation 121, 122, 126–130, 132, 138, 169
 - variable 127, 130, 133, 134, 136, 137, 141, 142
- Bernoulli 32
 - design 32
 - sampling 32, 51
- between-strata variance 66
- bias 17, 19
 - conditional 235
 - of the expansion estimator 19
 - of the Hájek estimator 359
- biased 19, 20, 81, 197
 - Hájek estimator 359
 - ratio estimator 197
 - sample 19
- Bienaym–Chebyshev inequality 298
- bootstrap 331
- bounded sequence 296
 - in probability 296
- bound management 253
- Brewer method 109
- burden 174
- c**
- calibration 10, 141, 190, 218–220, 222, 223, 226, 236–261, 277, 313, 315, 318, 321, 376
 - bound management 253
 - equation 238
 - estimator 190, 328
 - existence of a solution 254

- calibration (*contd.*)
 - generalized 256–258
 - general pseudo-distance 246
 - on households and individuals 255
 - improper calibration function 254
 - method
 - Deville 249
 - empirical likelihood 242
 - linear 239
 - logistic 249
 - raking ratio 240
 - reverse information 244
 - Roy and Vanheuverzwyn 251
 - truncated linear 245
 - in practice 258
 - regression estimator 229
 - solving algorithm 252
 - variable 242, 255, 257–259
 - central limit theorem 2
 - cluster 143
 - combined ratio estimator 199
 - compatibility of a sample 133
 - complex statistic 295
 - conditional expectation 24
 - construction of the strata 77
 - convergence in probability 296
 - convex polyhedron 129
 - coordination 172
 - corrected variance 13
 - correlation coefficient 14
 - cost function 134
 - Cotton and Hesse method 176
 - covariance 14, 15, 202, 283, 306–308, 332, 377
 - coverage 19, 333, 334
 - error 334
 - cube method 125
 - cumulative distribution function 284
 - curve of Lorenz 290, 292
- d**
- decomposition into simple random
 - sampling designs 104
 - design
 - based approach 1, 13
 - cluster 143
 - expectation 15
 - multi-stage 157
 - self-weighting 156
 - simple 209
 - two-phase 161
 - two-stage 148
 - with unequal probabilities 83–115
 - variance 15
 - Deville systematic sampling 91
 - difference estimator 195, 201, 202, 204–207
 - distance 134, 167, 169, 238
 - distinct units 23, 27, 38–41, 43–45, 88, 115
 - divergence of Kullback–Leibler 220
- e**
- eigenvector 32
 - empirical likelihood 7, 62
 - entropy 14, 58–61, 85, 93–95
 - Bernoulli design 58
 - maximum 112
 - for simple designs 58
 - for systematic designs 58, 91
 - enumerative statistics 2
 - equation
 - balancing 121, 122, 126–130, 132, 138, 169
 - calibration 238
 - estimating 310
 - error
 - due to nonresponse 334
 - of coverage 334
 - of measurement 334
 - of overcounting 334
 - sampling 333
 - of undercounting 334
 - estimating equation 310
 - estimation
 - calibration 221, 237–259
 - of complex parameters 281–293
 - conditional 235
 - of a covariance 283
 - of the cumulated income 288
 - cumulative distribution function 284
 - by expansion 18
 - of a function of totals 281
 - of a geometric mean 281

- of the Gini index 290
 - of an implicit function 283
 - of the Lorenz curve 289
 - of a mean 19
 - of a proportion (simple random sampling) 47
 - of a quantile 284
 - of the quintile share ratio 289
 - of a ratio 48, 281
 - simple random sampling 48
 - regression 236
 - multiple 225–236
 - in a stratified design 66
 - of a total 18
 - of the variance 282, 303
 - by linearization 295
 - in two-stage designs 153
 - unequal probabilities 114
 - estimator
 - calibrated 226, 260
 - by calibration 190, 217, 237, 328
 - generalized 257, 258
 - by difference 202
 - by expansion 7, 18–21, 23, 25, 81, 111, 166, 183–187, 195, 196, 200, 202, 205, 206, 208, 211, 212, 216, 217, 221, 223–225, 227, 234, 235, 237, 239, 259, 278–281, 359, 366
 - under balanced sampling 119, 120, 122, 123, 135, 137, 141
 - under Bernoulli sampling 34, 35
 - under cluster sampling 145–147
 - of the distinct units 38, 41, 42, 44, 45
 - under simple random sampling 28, 31
 - under stratification 67, 69–71, 74, 77
 - under systematic sampling 57
 - under two-stage sampling 151, 155, 157, 158, 162, 164
 - under unequal probability sampling 83, 93, 110–112
 - Hájek 145, 376
 - of Hansen–Hurwitz 22
 - of Horvitz–Thompson 18
 - of the mean 20, 31, 44, 80, 82, 211, 165, 166
 - by multiple regression 309, 328, 331, 377
 - post-stratified 211, 212
 - ratio 211, 231, 268, 303, 375, 376
 - regression 190, 224–226, 228, 230, 232, 233, 240, 255, 257, 275, 277, 301, 309, 318, 329, 373
 - generalized 226
 - multiple 225–231, 236, 307
 - optimal 226, 236, 258
 - of the total 31, 85, 159, 196, 211, 229, 238, 256, 366
 - of the variance 21
 - expansion estimator 7, 18–21, 23, 25, 28, 81, 111, 166, 183–187, 195, 196, 200, 202, 205, 206, 208, 211, 212, 216, 217, 221, 223–225, 227, 234, 235, 237, 239, 259, 278–281, 359, 366
 - under balanced sampling 119, 120, 122, 123, 135, 137, 141
 - under Bernoulli sampling 34, 35
 - under cluster sampling 145–147
 - of the distinct units 38, 41, 42, 44, 45
 - under simple random sampling 28, 31
 - under stratification 67, 69–71, 74, 77
 - under systematic sampling 57
 - under two-stage sampling 151, 155, 157, 158, 162, 164
 - under unequal probability sampling 83, 93, 110–112
 - expectation 15
 - conditional 24
 - under the design 15
 - of the estimator 17
 - under the model 263
 - of the sample size 16
 - experimental statistics 5
- f**
- finite population correction 31
 - flight phase 130
 - function
 - calibration 238–240, 242, 245, 246, 249–255, 257
 - improper 254
 - of influence 314–319

function (*contd.*)

of interest 17, 25, 38, 48, 65, 74, 75, 147,
217, 218, 235, 289, 301–303, 307,
310, 314–316, 318, 320, 321, 331

functional 314–318

g

generalized random tessellation stratified
(GRTS) 167

geographic stratification 65

geometric mean 281, 305

Gini index 290–292, 315, 320, 323, 330

g-weight 227, 228, 230, 259, 328

h

Hájek estimator 20, 42

Hansen–Hurwitz estimator 87

heteroscedasticity 231, 269, 273–277

heteroscedastic model 263, 267, 269

history of survey sampling 1

homoscedastic model 267

Horvitz–Thompson

estimator 18

variance estimator 22

i

implementation of the simple random
sampling designs 51

implicit function of interest 310

imputation 342

hot deck 346

by the nearest neighbor 342

by prediction 342

inclusion probability 15, 98, 99

for sampling with unequal probabilities
83

under simple random sampling 28

in stratified designs 68

independence 149

index

Gini 290–292, 315, 320, 323, 330

inequality 288, 295, 301

Moran 170

modified 172

indicator variable 6, 14, 24, 32, 185, 191

inequality

of Bienaym–Chebyshev 298

index 288, 295, 301

inferential statistics 13

influence function 314–319

information

auxiliary 2, 9, 10, 32, 65, 72, 77, 83, 111,
120, 143, 195, 196, 200, 201,

209–211, 217, 225, 226, 237,

255–257

reverse 244

interest function 147

invariance 149

invariant linearly 146, 147

inverse sampling with replacement 45

item nonresponse 133, 333, 334, 342

j

jackknife 331

joint inclusion probability 15, 98

k

Kish

individual 256

and Scott method 176

k nearest neighbors (kNN) 343

Kullback–Leibler divergence 220, 240, 248

l

label 13

landing phase 130, 133

by linear programming 133

by relaxing variables 135

law of total variance 5, 43, 66, 151, 211,
279

likelihood 242, 244, 312

empirical 242–245, 250, 254

linearization 295–331

of a calibrated estimator 313, 327

of the coefficient of a logistic regression
311

of a complex function with calibrated
weights 329

of a covariance 306

of a double sum 319

of an estimator by multiple regression
309

of a function of totals 301, 316

of a geometric mean 305, 326

- of the Gini index 320, 323, 330
- by the indicator variables 324
- of a mean 325
- of a multiple regression estimator 328
- of a parameter
 - of a calibration equation 313
 - defined by an implicit function 318
- of a quantile 326
- of a ratio 303, 325
- of a ratio estimator 304, 323
- of a regression coefficient 308
- by steps 308, 318
- of a total 316, 325
- of a univariate regression estimator 309
- of a variance 305, 321
- of a vector of regression coefficients 307
- linearized 295–331
- linearly invariant 20, 45, 146, 147
- local
 - cube method 169
 - pivotal method 109
- Lorenz curve 289, 290, 292
- m**
- mathematical statistics 1, 7
- maximum
 - entropy 58–61, 83, 95, 96, 98, 111, 112, 138
 - likelihood 7, 62
- mean
 - estimator 25, 31, 38, 70, 71, 360
 - geometric 281
 - of the population 325
 - sample 212
 - squared error 18, 49, 88, 365
- measure of spread 170
- method
 - calibration 237
 - Cotton and Hesse 176
 - Deville 249
 - of Deville, Grosbras and Roth 124
 - empirical likelihood 242
 - Kish and Scott 176
 - linear 239
 - truncated 245
 - local cube 169
 - local pivotal 169
 - logistic 249
 - Netherlands 178
 - raking ratio 240
 - reverse information 244
 - Rivière 177
 - Roy and Vanheuverzwyn 251
 - Swiss 178
 - travelling salesman 169
- microstrata 178
- minimal support 91, 103, 122, 123, 141
- minimum
 - entropy 61
 - support design 103
 - variance 38, 81, 277
- missing completely at random (MCAR) 335
- missing at random (MAR) 335
- missing not at random (MNAR) 335
- model
 - assisted approach 1, 9
 - based approach 1, 8, 263–280
 - heteroscedastic 267, 269
 - homoscedastic 267
 - superpopulation 263–280
- multinomial
 - sampling 86
 - vector 56
- multiple
 - frame survey 182
 - regression estimator 307, 309, 328, 331, 377
 - selection 24
 - sources 11, 182
- n**
- Narain procedure 89
- nearest neighbor 343
- negative estimator of variance 22
- nonresponse 333
 - and calibration 339
 - completely at random 335
 - confounded 335
 - item 133, 333, 334, 342
 - at random 335
 - questionnaire 333, 334, 336, 341, 348
 - uniform 335

O

observation unit 13
 official statistics 2, 3, 9
 optimal
 allocation 6, 71
 stratification 6, 71, 74
 optimality 8, 65, 75, 76
 and cost in stratification 76
 in stratification 74
 order
 of magnitude 295, 296
 in probability 295, 296
 stochastic 296
 pivotal method 108
 sampling 100

P

parameter 13, 17
 estimation 17
 permanent random numbers (PRNs) 172,
 175–178
 pivotal
 local method 169
 method 107
 Poisson
 design 92
 sampling 92–95, 100, 163, 165, 179
 population 13
 covariance 14
 mean 13, 66, 119, 325
 size 11, 13, 19, 51, 53, 56, 65, 73, 79, 86,
 104, 105, 108, 119, 123, 134, 207,
 363, 364, 366, 376
 stratified 271
 total 13
 post-stratification 209, 210, 217,
 221, 255
 post-stratum 209–212, 218, 231
 probability
 inclusion
 in cluster designs 145
 two-stage designs 150
 procedure
 of Rao–Sampford 98
 splitting 102
 proportion 47, 48, 62, 70, 182, 288, 289,
 362, 363

proportional
 allocation 70
 stratification 70
 pseudo-distance 246, 260
 general 246
 purposive selection 5

Q

quality of balancing 135
 quantile 284
 questionnaire nonresponse 333, 334, 336,
 341, 348
 quintile share ratio 289

R

raking ratio 224, 240–242, 246
 random
 pivotal method 109
 sample 8, 14–17, 22, 25, 34, 38, 52,
 54–56, 66–68, 115, 119, 125, 138,
 200, 237
 sample size 85, 96
 selection 5
 sorting method 52
 vector 14
 Rao–Blackwellization 23
 Rao–Blackwellized estimator 42
 Rao–Blackwell theorem 44
 Rao–Sampford rejective procedure 98
 ratio 2, 14, 20, 48, 49, 195–200, 202, 204,
 206–208, 211, 281, 303, 373, 375
 estimator 2, 195–200, 202, 204,
 206–208, 304, 321, 323, 373, 377
 bias 197
 raking 218
 regression
 estimator 190, 195, 202–205, 207, 208,
 301, 309, 318, 328
 multiple 225–236
 univariate 309
 rejective 98
 method 119
 representative
 method 4
 sample 72
 stratification 70
 representativeness 4–6

- reservoir method 54
 - retaining of the distinct units 38
 - reverse approach 152, 154, 165, 347–349
 - reweighting 11, 336, 342, 343, 347, 357
 - Rivière method 177
 - robust 9, 236, 273, 277, 314
 - statistic 314
 - robustness 9, 269, 314
- S**
- salesman travelling 169
 - Sampford algorithm 100
 - sample 14
 - balanced 271, 275
 - independent samples 163
 - mean 69, 70, 119, 369
 - random 144, 149, 158, 161, 217, 221, 269, 278, 280
 - size 7, 15–17, 20–23, 26, 27, 29, 32–36, 38, 41, 44, 49–51, 57, 59–62, 65, 68, 71, 73, 76, 77, 79–81, 83, 85, 86, 88, 89, 92, 94–98, 100, 112, 113, 116, 117, 119, 121, 122, 124, 126, 134, 135, 137, 139, 141, 145, 147, 156, 157, 199, 205, 272, 361, 363, 364, 368, 370, 371
 - random 85, 96
 - sampling
 - cluster 143
 - with equal probability 146
 - with unequal probability 147
 - design 14
 - with equal probability 155
 - error 333
 - repeated 172
 - with replacement 22
 - of distinct units 38
 - with unequal inclusion probabilities 86
 - with unequal probabilities 83–115
 - second-order inclusion probability 15
 - selection-rejection method 53
 - sequence
 - bounded 296
 - in probability 296
 - of lower order 296
 - in probability 296
 - sequential 51, 56, 86, 100, 108
 - algorithm 56, 108
 - pivotal method 108
 - sharing of weights 188, 190, 191
 - simple
 - design 27–57
 - random
 - sampling without replacement 27
 - sampling with replacement 36, 56
 - size
 - of the cluster 143
 - population 11, 13, 19, 51, 53, 56, 65, 73, 79, 86, 104, 105, 108, 119, 123, 134, 207, 363, 364, 366, 376
 - of post-stratum 209
 - of the sample 7, 15–17, 20–23, 26, 27, 29, 32–36, 38, 41, 44, 49–51, 57, 59–62, 65, 68, 71, 73, 76, 77, 79–81, 83, 85, 86, 88, 89, 92, 94–98, 100, 112, 113, 116, 117, 119, 121, 122, 124, 126, 134, 135, 137, 139, 141, 145, 147, 156, 157, 199, 205, 272, 361, 363, 364, 368, 370, 371
 - random 85, 96
 - of the stratum 65
 - spatial sampling 167
 - splitting method 101
 - spread 170
 - statistic 15, 17
 - asymptotic 295
 - complex 295
 - robust 314
 - statistical unit 175, 183
 - statistics 1
 - mathematical 1
 - official 9, 157, 226
 - Stirling number of the second kind 39
 - stratification 58, 65–79, 111, 141, 143, 149, 167, 366
 - optimal 366
 - variable 71, 77, 79
 - stratified design 65–79
 - with optimal allocation 71, 74
 - with proportional allocation 70
 - stratum 65–82, 121, 172, 176, 200, 236, 255, 272, 366, 368, 369, 375, 376
 - total estimator 69

successive drawing 88
 of the units 52
 sufficient statistic 88
 sum
 of the inclusion probabilities 16, 17,
 134, 170
 of the variances 68, 79
 of weights 359
 superpopulation 8, 9, 263–280, 347
 model 8, 9, 263
 survey
 cluster 143
 two-stage 148
 Swiss method 178
 systematic 89
 sampling 57–58
 of Deville 91
 with equal probabilities 27, 57
 with unequal probabilities 58,
 89, 90
 unequal probabilities 167

t

take-all stratum 79
 tessellation 167
 total
 estimation 18
 estimator 20, 58, 65, 69, 70, 77, 80, 183,
 190, 191, 313, 331, 338,
 339, 348
 variance 5, 43, 66
 weight 80
 travelling salesman 169
 two
 phases 161
 design 336
 stages 148

u

unbiased 7, 17, 20, 22, 24, 25, 87, 156, 159,
 160, 162, 199, 201, 205, 214, 215,
 226, 230, 231, 264, 265, 277–280,
 360, 373
 approximately 203, 237
 conditionally 42, 45
 for designs with replacement 38

estimation 25, 87, 146, 147, 158, 163,
 218
 estimator 7, 25, 31, 62, 81, 140, 141,
 154, 227, 235, 236, 266
 under balanced sampling 141
 BLU 264
 class of 8
 condition for 19
 definition 17
 of a double sum 21
 of Hansen–Hurwitz 23, 36
 of Horvitz–Thompson 114
 of Sen–Yates–Grundy 22
 under simple random sampling
 31
 in a stratum 67
 of variance 23, 25, 62, 154
 unconditionally 42
 unequal probability
 Brewer method 109
 Deville’s systematic sampling 91
 maximum entropy sampling 95
 minimum support design 103
 order sampling 100
 pivotal method 107
 Poisson sampling 92
 Rao–Sampford sampling 98
 sampling 7, 83–115
 approximation of the variances
 111
 with replacement 6, 86
 splitting method 101
 successive drawing 88
 systematic sampling 7, 89
 variance estimation 114
 unit
 primary 149
 secondary 149
 statistical 175, 242

v

variable
 auxiliary 229, 261, 331
 balancing 130, 133, 134, 136, 137, 141,
 142, 172
 calibration 242, 255, 257–259

- indicator 209, 210, 231, 242
 - of interest 8, 9, 20, 32, 34, 58, 65, 71, 72, 74, 79, 111, 190, 191, 200, 202, 217, 225, 226, 236, 258, 263, 284, 285, 316
 - variance 15
 - approximation 137
 - covariance 16, 24, 26, 313
 - estimation 140
 - estimator 21
 - inter-primary unit 152
 - intra-primary unit 152
 - under the model 263
 - total 151, 211, 279
 - of the total estimator 20
 - in two-stage designs 151
 - vector
 - of dilated values 20
 - of inclusion probabilities 15, 97, 103, 104
 - multinomial 56
 - of ones 16
 - random 14, 22
 - of target probabilities 100
 - Voronoi polygon 170
- W**
- weight 19, 80, 138, 187, 190, 196, 211, 212, 218, 221, 226–228, 230, 234, 237–246, 248, 249, 251–261, 264, 277–279, 281, 283–286, 289, 291, 292, 301, 315, 322, 324, 327–330, 338–340, 343, 344, 348, 376
 - bounded 246, 253
 - choice 231
 - distortion 228
 - extreme 250
 - g 227
 - positive 249, 251
 - of the sampling frame 186
 - sharing 188, 190, 191
 - of a survey 178
 - of units 191
 - weighting 226, 237, 239, 281, 336, 342, 343, 347, 357
 - within-strata variance 66
- y**
- Yates–Grundy
 - variance 21
 - variance estimator 22