

Index

- α and β errors, 151
- a priori weighting, 48
- accuracy, 3, 80, 134
- AID, 200
- alternative hypothesis, 148
- Automatic Interaction Detection, 200
- bootstrap, 86, 88
- calibrated confidence limits, 123
- calibrated error, 109
- calibrated sample size
 - combined, 108
 - definition, 105
 - properties, 107
- canonical correlations, 182
- categorical variables, 172, 176, 181, 186, 193, 200, 201, 206
- Cauchy–Schwarz inequality, 245
- Central Limit Theorem, 142, 232
- CHAID, 200
- chi-square distribution, 166, 169
- chi-square test
 - degrees of freedom, 167, 169
 - of fit, 165
 - of independence, 168
- classification and regression trees, 200
- cluster analysis, 192, 194, 206
- cluster sampling, 24, 77
- clustering, 82, 84, 95
- clustering effect, 111
 - spurious, 97
- complex samples, 25
- confidence limits, 233
- contingency tables, 172, 178
- correlation coefficient, 175, 235
- correspondence analysis, 186
- covariance, 235
- C&RT, 200
- data fusion, 179, 215
- density function, 229
- design effect, 5, 19, 76, 101, 111, 167, 169
 - and clustering, 95
 - and stratification, 92
 - combined, 94
 - definition, 90
 - in a drug survey, 116
 - in data analysis, 99
 - properties, 91
- disproportionate stratification, 19
- distribution
 - beta, 240
 - beta-binomial, 239
 - binomial, 239
 - chi-square, 240
 - double exponential, 240
 - exponential, 240
 - gamma, 240
 - geometric, 239
 - hypergeometric, 239
 - logistic, 240
 - negative binomial, 239
 - normal, 229, 240
 - Poisson, 239
 - uniform, 240
- donor (data fusion), 216
- effective sample size, 5, 232
 - definition, 90
 - properties, 91
- entropy, 184, 205, 210
- factor analysis, 192
- finite population correction, 43

- histogram, 228
- hot-deck imputation, 133, 216
- hypothesis, 144
- imputation, 132
- imputation error, 136
- indexed values, 178, 187, 221
- intra-class correlation, 95
- jackknife, 86, 87
- Keyfitz, 86
- Kolmogorov–Smirnov test, 170
- Lagrange multipliers method, 241
- missing data, 132
 - imputation techniques, 132
 - variance estimates, 134
- multivariate techniques, 192
- non-response, 10, 32, 77
- non-response weighting, 48
- null distribution, 149
- null hypothesis, 145, 146, 148
 - of ‘no difference’, 146, 148, 159
- numerical and categorical variables, 172
- one-tailed test, 153
- P*-value, 149
- population, 10, 42, 227
 - small, 42
- population registers, 11
- post-stratification, 22
- practical significance, 146
- precision, 3, 81, 92, 109
- principal component analysis, 192
- proportionate stratification, 18
- quota sampling, 19, 36, 88
- ratio estimates, 82, 84
- recipient (data fusion), 216
- regression analysis, 197
- regression to the mean, 221
- response rate, 32, 33, 35, 42, 136
- sample
 - representative, 8, 27, 227
 - simple random, 16
 - unbiased, 17, 227
- sampling
 - constraints, 9
 - objectives, 5
 - plan, 6
 - priorities, 7
 - probability, 27
 - with unequal probability, 25
 - sampling frame, 11
 - sampling method
 - cluster sampling, 24
 - multi-stage, 17
 - non-probability, 8, 41, 88
 - convenience sampling, 41
 - judgement sampling, 42
 - purposive sampling, 42
 - quota sampling, 36
 - probability, 8
 - proportional to size, 27
 - stratification, 18, 81
 - disproportionate, 18, 19
 - optimised allocation, 21, 92
 - systematic sampling, 23
 - sign test
 - for difference in rating
 - for median
 - significance level, 145, 149
 - significance testing
 - choice of a test, 151, 155
 - effect of weighting, 154
 - meaning of significance
 - principles, 151
 - purpose and philosophy, 140
 - statistical packages, 155
 - with non-simple-random samples, 153
 - with weighted data, 154
 - software, statistical, 125, 155, 198
 - standard deviation, 231
 - standard error, 5, 90, 109, 231
 - statistic
 - chi-square, 169
 - Kolmogorov–Smirnov, 170
 - statistical significance, 146
 - stratification, 18, 77, 81, 84, 109
 - stratification component
 - of design effect, 99, 117
 - of weighting effect, 108, 117
 - stratification effect, 19
 - strength of association, 181
 - Cramer’s coefficient, 181
 - entropy coefficients, 184
 - predictive coefficients, 182
 - strength of fit, 178
 - chi-square distance, 180
 - Euclidean distance, 180
 - geometric distance, 180
 - surrogate lists, 12
 - surrogate procedures, 15

- t*-test, 157
- test for
 - a single value, 157, 158
 - correlated proportions, 162
 - goodness of fit
 - homogeneity, 168
 - independence, 168
 - independent means, 159
 - independent proportions, 161
 - repeated measurements, 164
- two-tailed test, 153
- type I and type II errors, 151

- universe, 227

- variance estimates, 79
 - clustering, 82
 - if there are missing data, 134
 - jackknife and bootstrap, 86
 - Keyfitz method, 86
 - post-stratification, 82
 - quota sampling, 88
 - ratio estimates, 83
 - ratio of ratios, 83
 - resampling, 87
 - stratification, 81
 - stratified clustering, 84, 85
 - stratified ratios, 84
 - systematic sampling, 85

- weighting
 - definition, 45
 - multi-stage and hybrid, 74
 - principles, 45
 - statistical reasons, 119
 - statistical software, 125
 - usage, 47
 - utility, 120
- weighting effect, 76, 109, 111
 - and design effect, 106, 109, 116
 - combined, 108
 - definition, 105
 - in a drug survey, 116
 - maximum value, 107
 - properties, 107
- weighting methods
 - aggregate weighting, 66
 - GREG, 73
 - iterative, 71
 - quantity weights, 73
 - variance minimisation, 73
 - cell weighting, 52
 - marginal weighting, 55
 - GREG, 65
 - iterative, 59
 - variance minimisation, 63
- weights, replication, 126

- Z-test, 157, 158, 159, 162