

Index

- agriculture, 19, 26, 29, 33, 60, 77, 107, 119, 125, 168; *see also* bovine tuberculosis, foot-and-mouth disease, field trial
- AIDS, 3, 61, 71, 166
- algorithmic methods, 10, 185
- analysis
 - change of, 17
 - of variance, 42, 177
 - plan of, 16
 - preliminary, 75–89, 178
 - principal components, 63
 - subgroup, 169
 - tabular, 86–87
 - unit of, 5, 18–20, 114
- anecdotal evidence, 36
- antithetic variable, 113
- Armitage–Doll model, 99
- atomic null hypothesis, 145
- attenuation, 70
- autocorrelation, 178
- autoregressive process, 132
- baseline variable, 161
- Bayes, Thomas, 143
- Bayes factor, 131, 150
- behaviour
 - animal, 54
 - anomalous, 79, 126
 - electoral, 35
 - human, 54
 - individual, 167
 - of a system, 55
 - political, 181
 - wildlife, 2, 17
- Berkson, J., 167
- Berkson error, 72
- bias, 21
 - ecological, 19
 - publication, 180
- blinding, 21, 40
- blood pressure, 1, 6, 38, 70, 72, 94, 136
- body mass index (BMI), 47, 62, 94, 172
- Bonferroni correction, 154
- bootstrap, 149
- bovine tuberculosis, 2, 17, 39, 77, 108, 114, 135
- cancer, 4, 48, 50, 73, 96, 99, 154, 160, 167, 172
- capture–recapture sampling, 36
- case–control study, 49, 73, 135
- causality, 122, 160–167
- censoring, 61–62
- choice-based sampling, 49
- climatology, 2
- clinical trial, 18, 25, 106, 125, 154, 168, 181
- cluster
 - detection, 81
 - randomization, 38, 114
- Cochran, W. G., 51
- component of variance, 25, 126
- concealment, 7, 21, 40
- confidence
 - distribution, 141
 - limits, 141
- Consumer Prices Index, 64
- contingency tables, 4, 175
- contrasts, 41, 93, 124, 125, 140, 173
- cost, 8, 11, 25, 28, 33, 53
- cross-classification, 43, 123

- cross-over experiment, 18, 114
- cross-sectional study, 46, 164
- data
 - large amounts of, 3, 10, 49, 53, 184
 - longitudinal, 48, 170, 179
 - missing, 9, 80
 - use of in statistical theory, 12
- data audit, 9, 76–77
- data-generating model, 169
- data mining, 3, 184
- data quality, 3
- data screening, 9, 77
- decision, 2
- delta method, 142
- derived variable, 62
- design, 5
 - and analysis, relationship between, 16
 - of investigation, 2
 - randomized block, 34, 40
 - unbalanced, 127
- diabetes, 47, 58, 102, 105, 164
- diet, 19, 72, 133, 174
- dimensionality, 8
- distribution
 - confidence, 141
 - exponential, 122
 - gamma, 122
 - log normal, 57, 96, 109, 122
 - multinomial, 98
 - normal, 26
 - Poisson, 102, 109, 142, 152
 - posterior, 142, 149, 157
 - prior, 143, 150, 157
 - Weibull, 122, 130
- dividing null hypothesis, 148
- domain of study, 33
- ecological bias, 19
- economics, 8, 11, 25, 28, 62, 87, 106, 165
- effect reversal, 171
- empirical Bayesian argument, 144, 156
- error, 93
 - Berkson, 72
 - measurement, 1
 - random, 24
 - systematic, 15, 70
- ethics, 75
- event history, 48
- experiment
 - definition of, 6
 - natural, 165
- exponential distribution, 122
- extensive variable, 57, 119
- extreme events, 37, 108
- factorial principle, 15, 26, 45
- false discovery rate, 156
- field trial, 17, 19, 28, 29, 33, 60, 77, 107, 119, 125, 153, 168
- Fisher, R. A., 165, 167
- fixed effects, 125
- fluid dynamics, 120, 152
- foot-and-mouth disease (FMD), 182
- forest plot, 85
- frequency domain, 179
- frogs, toe clipping of, 36
- gamma distribution, 122
- gender, 58, 94, 102, 133, 145, 168, 174, 177, 186
 - as a possible cause, 161
- generalized least-squares, 128
- genome-wide association study (GWAS), 152
- geographic information system (GIS), 82
- gold standard diagnosis or method, 68, 69
- graphical methods, 10, 82
- gross domestic product (GDP), 8
- growth curves, 56, 111
- Hawthorne effect, 55
- hierarchical structure, 112, 116, 131, 157
- HIV, 3, 40, 62, 71, 166
- hormone replacement therapy (HRT), 7, 172
- hypertension, 6
- hypothesis, 2
 - null, 25, 60, 141, 145
 - atomic, 145
 - dividing, 148
- ideal sequence, 2, 188
- individual investigator, role of, 5
- intensive variable, 57
- intention to treat, 39
- interaction, 44, 111, 125, 129, 132, 138, 171, 186
- intermediate variable, 60, 161
- interpretation, 2, 104, 159
- intrinsic variable, 59, 102

- inverse probability, 143
- judgement, 21, 25, 40, 96, 107, 126, 186
- judicial sentencing, 116
- King, David, 182
- latent
 - class model, 68
 - variable, 64
- Latin square, 43
- least-squares method, 10, 16, 22, 38, 42, 70, 126, 133, 185
 - generalized, 128
- length-biased sampling, 32
- likelihood, 130, 151
- log normal distribution, 57, 96, 109, 122
- logistic regression, 49, 80, 87, 116, 125, 135, 185
- longitudinal data, 48, 170, 179
- machine learning, 185
- manifest variable, 64
- Markov chain Monte Carlo (MCMC), 149
- matching, 17, 37, 49, 51
- maximum likelihood, 130, 151
- measurement error, 1, 69
- measurements
 - cost of, 8
 - dimensionality of, 8
- meta-analysis, 2, 15
- Michaelis–Menten equation, 100
- missing data, 9, 80
- model
 - Armitage–Doll, 99
 - data-generating, 169
 - latent class, 68
 - multi-level, 180
 - nonparametric, 95
 - one-hit, 99
 - parametric, 95
 - probability, 10, 92–104
 - quasi-realistic, 92
 - semiparametric, 95
 - stochastic, 92, 100
 - toy, 92
 - two-hit, 99
- monetary-unit sampling, 30
- multi-level model, 180
- multi-stage sampling, 34
- multinomial distribution, 98
- multiple testing, 155
- multivariate
 - normal distribution, 66
 - outlier, 79
- natural
 - experiment, 165
 - variability, 1, 93
- nesting, 44
- Neyman, J., 167
- noise factors, 168
- noncompliance, 39, 106
- nonlinearity, 138
- nonparametric model, 95
- nonspecific variable, 125
- normal distribution, 26
 - multivariate, 66
- nuisance parameter, 97, 118, 143
- null hypothesis, 25, 60, 141, 145
 - atomic, 145
 - dividing, 148
- observational study, 20, 164
 - definition of, 7
- Occam’s razor, misuse of, 151
- one-hit model, 99
- outlier, 78
 - multivariate, 79
 - spatial, 82
- overview, 2, 15, 130
- panel of individuals, study of, 35
- parameter
 - nuisance, 97, 118, 143
 - of interest, 90, 118
- parametric model, 10, 95
- performance target, 55
- philosophy, 13, 147, 151
- physics, 57, 64, 88, 91, 142, 147, 152, 166, 171, 184
- pilot study, 25, 76
- plan of analysis, 16
- Poisson
 - distribution, 102, 109, 142, 152
 - process, 88, 99, 122
- policy, 63, 108, 160
 - implications for, 47
 - presentation of results which inform, 181–183
- polynomial regression, 131
- population attributable fraction, 120
- posterior distribution, 142, 149, 157
- power, 25

- precision, 3, 15, 23, 32, 53, 81, 112, 119, 126, 129, 151, 173
 prediction, 11
 preliminary analysis, 75–89, 178
 pressure
 of reactants, 136
 of time, 75
 to produce one-dimensional summaries, 9
 to produce particular results, 181
 to stick to plan of analysis, 186
 principal components analysis, 63
 prior distribution, 143, 150, 157
 probability, 104
 model, 10, 92–104
 proportional adjustment, 33
 publication bias, 180
 quality control, 4
 quality of life, 8, 65, 81
 quasi-realistic model, 92
 questionnaire, 54, 57, 65, 80, 135
 random
 effects, 125
 error, 24
 variable, 94
 randomization, 18
 cluster, 38, 114
 randomized block design, 34, 40, 126
 range of validity, 15, 27, 108
 recall bias, 46, 49, 162
 recovery of inter-block information, 128
 regression, 33, 63, 65, 69, 95, 110, 121, 136, 138, 150, 155, 163
 chain of, 59, 102, 170
 dilution, 70
 logistic, 49, 80, 87, 116, 125, 135, 185
 polynomial, 131
 regulatory agency, 16
 inhibitory role of, 5
 relative risk, 58, 85, 172
 reliability, measurements of, 8
 replication, 15
 research question, formulation of, 2
 residual, 41
 respondent-driven sampling, 36
 response
 curve, 20, 56
 surface, 46, 136, 173
 Retail Prices Index, 64
 Reynolds number, 121
 robust methods, 79
 routine testing, 4
 sampling
 capture–recapture, 36
 choice-based, 49
 frame, 30
 inspection, 9
 length-biased, 32
 monetary-unit, 30
 multi-stage, 34
 respondent-driven, 36
 snowball, 36
 systematic, 30
 temporal, 34
 scale of effort, 15
 semiparametric model, 95
 sensitivity, 70
 sequence, ideal, 2, 188
 serial correlation, 178
 Severe Acute Respiratory Syndrome (SARS), 92
 significance test, 25, 145–156
 simulation, 92, 104, 113, 131, 146, 148
 smoking, 120, 137, 160, 167, 172
 smoothing, 10, 85, 96
 snowball sampling, 36
 social mobility, 176
Sordaria, 98, 105
 sparsity, 138
 spatial outlier, 82
 specificity, 70, 168
 split-plot design, 19
 standing, equal, 170
 stochastic model, 92, 100
 study
 case-control, 49, 73, 135
 cross-sectional, 46, 164
 observational, 7, 20
 pilot, 25, 76
 stylized facts, 109
 subgroup analysis, 169
 sum of squares, 42, 150
 surrogate
 endpoint, 60
 variable, 2
 survival analysis, 61, 96
 synthesis of information, 2
 systematic
 error, 15, 21, 70
 sampling, 30
 variation, 109

- tabular analysis, 86–87
- target, performance, 55
- temporal sampling, 34
- textile, 16, 23, 38, 163
- theories, elaborate, 165
- time series, 11, 63, 104, 106, 109, 132, 178
- time scale of investigation, 16, 187
- tolerance, 66
- toy model, 92
- transformation, 119, 129, 136, 174
- transmission of infection, 40, 82, 92, 100; *see also* bovine tuberculosis, foot-and-mouth disease, HIV
- transparency of analysis, 9
- trial
 - clinical, 18, 25, 106, 125, 154, 168, 181
 - field, 17, 19, 28, 29, 33, 60, 77, 107, 119, 125, 153, 168
- two-hit model, 99
- unbalanced design, 127
- unbiased estimate, 71
- unit of analysis, 5, 18–20, 114
- validity, 8, 53
 - range of, 15, 27, 108
- variability, natural, 1, 93
- variable
 - antithetic, 113
 - baseline, 161
 - derived, 62
 - extensive, 57, 119
 - intensive, 57
 - intermediate, 60, 161
 - intrinsic, 59, 102
 - latent, 64
 - manifest, 64
 - nonspecific, 125
 - random, 94
- variation, systematic, 109
- visualization, 10
- Weibull distribution, 122, 130
- Yates, F., 51