

---

## Index

- Aalen's linear model, 179  
 absolute risk, 102, 182, 208  
 additive hazard model, 178  
 adjustment, 16, 58, 64, 72, 84, 92, 108, 113, 119, 186, 197, 235, 244  
   age-adjustment, 126  
 air pollution, 124  
 alcohol, 202  
 asbestos, 164, 186  
 asymptotic  
   covariance, 40  
   mean, 38  
   methods, 38, 60, 241  
   normality, 39, 40, 260  
   relative efficiency, 101, 130, 168, 202, 203, 210  
   standard error, 39, 43, 48  
   variance, 38–41, 43–45, 48, 50, 77, 100, 123, 257  
 asymptotically unbiased estimate, 38  
 attenuation, 6, 131, 215, 220, 222, 224, 226, 237, 238, 248  
 attributable  
   fraction, 109  
   risk, 102, 103, 109  
  
 balanced sampling design, 149  
 bankruptcy study, 28  
 baseline hazard estimator, 189  
 Bayesian methods, 89, 131, 238, 258  
 Berkson error, 222, 224, 237  
 bias, 5, 23, 24, 51, 73, 82, 109, 127, 153, 231, 238  
   checks for absence of, 27  
 biased design, 29, 175, 176, 188  
 binomial distribution, 38, 60, 68, 129  
 biobank, 149  
 bivariate normal distribution, 107  
  
 bootstrapping in case-cohort studies, 210, 217, 230  
 borrowing strength from collection of studies, 259  
 bovine TB, 28  
 brain cancer and exposure to pesticides, 29  
 breast cancer studies, 178, 243  
  
 canonical parameter, 60  
 career choice determinants, 27  
 case, definition of, 20  
 case-base study, 16, 193, 209  
 case-cohort study, *see* case-subcohort study  
 case-control study, advantages of, 11  
 case-crossover study, 105, 124, 125, 131  
 case-only study, 111, 122, 123, 131  
 case-parent study, 154  
 case-pseudosib study, 154  
 case-sibling design, 153  
 case-subcohort study, 18, 20, 161, 182, 191, 193, 209, 230  
   stratified, 204  
 causality, 14, 57, 58, 61  
 central limit theorem, 254  
 chi-squared  
   statistic, 50  
   test, 34, 242  
 choice-based sampling, 26, 28  
 classical measurement error, 222  
 Cochran's formula, 94  
 cohort study, 2, 8, 20–22, 32, 57, 81, 109, 161, 162, 165, 168, 178, 179, 181, 189, 192, 199, 251  
 coincidence, 129  
 collapsibility, 51, *see also* non-collapsibility

- colorectal cancer study, 65, 69, 78, 92, 185, 226
- combination of estimates, 14, 48, 51, 56, 61, 73, 136, 184, 197, 234, 241
- composite likelihood, 257
- concordant pair, 15, 66, 68, 70, 82, 118
- conditional  
 inference, 63, 91  
 logistic regression, 66, 92, 154  
 odds ratio, 4, 15, 48, 49, 52–54, 139
- confidence interval, odds ratio for, 38, 68
- confounder, 14, 59, 84, 114, 134
- continuity correction, 50
- contrast, 40
- controls, selection of, 22, 81
- Cornfield's inequalities, 59
- counter-matching, 151, 171, 173, 205
- cross-product ratio same as odds ratio, 35
- cross-sectional study, 2
- cumulative  
 hazard, 182, 183, 189, 208  
 log odds, 121  
 odds ratio, 131  
 sampling, 16
- curvature of the log likelihood, 255
- delta method, *see* local linearization
- diagnostic testing, 214
- diet, 65, 78, 92, 185, 226, 243
- difference of probabilities, 45, 102
- differential  
 error, 213, 226  
 recall, 25
- directed acyclic graph, 2, 7
- discordant pair, 15, 66, 68, 70, 82, 118
- econometrics, 26, 28, 30
- effect modification, 58
- efficiency, 100, 130, 134  
 alternative analyses of, 116  
 nested case-control study of, 168  
 of matched design, 82  
 of sampling, 151, 170
- EM algorithm, 239
- empirical Bayesian approach, 259
- employment discrimination, 79
- EPIC-Norfolk study, 65, 92, 185, 203, 226, 243
- etiologic fraction, 109
- event time, uncertainty in, 176
- exact method, 38
- excess risk, 179
- expected information, 253  
 in nested case-control study, 168
- experiment, randomized, 2, 62, 81
- explanatory variable, multilevel, 47
- exponential distribution, 60, 182
- family studies, 16, 152
- family-based design, 158
- Fisher information, 77, 253
- flat prior, 89, 259
- floating estimate, 48, 61
- frequency matching, 14
- gene–environment interaction, 124, 131, 133, 149, 153, 154, 158, 159, 173, 188, 237
- generalized  
 hypergeometric distribution, 37, 42, 43, 49, 60, 73, 249  
 least squares, 122
- genetic admixture, 153
- genetic epidemiology, 123, 153
- graphical models, 2, 61
- hazard rate, 54, 81, 82, 127, 161, 165
- hazard ratio, 81, 165, 194, 208
- heart disease, 26, 112, 178
- Horvitz–Thompson estimator, 258
- hospital control groups, 112
- household cleaning products and breast cancer, 214
- hybrid epidemiologic design, 193
- hypergeometric distribution, 37, 60
- incidence density sampling, 20, 54, 56, 61, 80, 82
- incidence rate, 54
- inconsistency of maximum likelihood, 75
- individual as own control, 125
- individual patient data, 247, 248, 251
- industrial machine defects, 129
- influence of individual on estimate, 201
- influence function, 210
- information, observed, 43, 69, 255
- information bias, 25
- information matrix, 88, 90, 116, 145, 148, 254
- interaction of variables, 3, 5, 58, 123, 134, 148, 149, 188, 205
- intrinsic variable, 64

- inverse  
 model, 42, 43, 46, 67, 79, 83, 88, 113  
 probability weighting, 206  
 weighting, 158
- ionizing radiation and breast cancer, 173
- Kronecker product, 218
- Lagrange multipliers, 88, 144, 145, 147
- Laplace transform, 107
- large-samples, *see also* asymptotic  
 properties, 116  
 theory, 69  
 limitations of, 256  
 values, 38
- latent variable, 121
- linear hazards model, 179
- linear odds model, 97, 98
- local linearization, 39, 217
- local odds, 106
- log odds ratio, *lods*, 59  
 asymptotic variance of, 41, 47, 70, 75,  
 77  
 preferred to odds ratio, 38
- log Poisson regression, 39
- logistic  
 difference, alternatives to, 46  
 distribution function, 43  
 function, definition of, 85  
 logistic regression, 13, 30, 79, 82, 88,  
 130  
 conditional, 66, 154
- lung cancer studies, 26, 27, 59, 97, 164,  
 186, 194
- Mantel–Haenszel estimate, 61, 109, 131
- marginal  
 distribution of exposure, information  
 in, 89, 110  
 odds ratio, 53
- matched study, 10, 15, 26, 30, 62, 81, 85,  
 91, 100, 109, 117, 132, 162, 169  
 analysis of, 65  
 misclassification in, 220
- maximum likelihood, 31, 37, 43, 50, 69,  
 72, 75, 76, 88, 116, 145, 147, 148,  
 183, 218, 246, 252, 254–256, 258  
 inconsistency of, 75
- measurement error, 25, 212
- melanoma and sunlamp use, 27, 41
- meta-analysis, 240
- misclassification, 212, 236
- missing data, 184, 189, 208, 211, 233
- mobile telephone use and minor traffic  
 accidents, 125
- moment reconstruction, 233, 234, 238
- multi-stage design, 158
- multinomial distribution, 39, 111
- multiple controls  
 design aspects of, 44  
 matched study in, 76
- multiple imputation, 184, 185, 208, 233,  
 238
- multiple regression, measurement error  
 in, 223
- multiplicative error, 225
- multiplicative intercept model, 97
- naive estimate, 228
- nested case-control  
 design, 191  
 sampling, 57  
 study, 20, 57, 81, 161, 191
- Neyman–Pearson theory, 60
- Neyman–Scott problem, 82
- non-collapsibility, 4, 52, 53, 61, 92, 93,  
 95, 100, 109, 120
- non-logistic model, 88, 109
- nondifferential error, 217
- nonparametric random effects, 72
- normal approximation to logistic, 229
- number of controls available, 76
- nutritional epidemiology, 237
- observed information, 43, 69, 255
- odds, 34
- odds ratio, 12, 17, 34, 37, 67  
 alternatives to, 46  
 empirical estimate, 38  
 same both ways, 12, 35, 84
- ordered logistic model, 97
- ordinal  
 data, 131  
 outcomes, 120  
 variable, treated as quantitative, 121
- outcomes, multiple, 111
- overmatching, 74, 81
- pairwise analysis, 116
- partial likelihood, 166
- path diagram, 1, 10
- pesticides and brain cancer, 29

- point process, 128, 129  
 Poisson  
   distribution, 39, 60, 85  
   process, 85, 90, 108, 126  
 political activism and career choice, 27  
 political science and case-control studies,  
   26, 28, 30  
 polychotomous  
   case-control studies, 111  
   response, 130  
 pool-adjacent violators, 121, 131  
 population  
   controls, 112  
   distribution, 32  
   model, 42, 67, 113  
   stratification, 153  
 population-based study, 21, 27, 161, 213  
 posterior distribution, 258  
 power calculations, 45, 82  
 primary base, 21  
 prior distribution, 258  
 proband, 155  
 profile likelihood, 108  
 proportional hazards model, 127, 165,  
   199, 210  
 proportional odds model, 97  
 prospective study, 2, 8, 21, 22, 32, 57,  
   161, 162, 165, 168, 178, 181, 189,  
   192, 199  
 prostate cancer risk factors, 157  
 pseudo-partial likelihood, 187, 188,  
   199–201, 204, 206, 207, 210  
 pseudo-likelihood, 88, 131, 144, 148,  
   179, 188, 201, 207, 211, 256, 257  
 pseudo-maximum likelihood, 210  
 pseudo-score function, 201, 206  
 publication bias, 244  
  
 Quebec asbestos miners, 164  
 quota sampling, 152, 174  
  
 radiation risks, 178, 188  
 radon exposure, 97, 224  
 random  
   digit dialling, 22  
   effects, 72, 131, 156, 245  
 randomization, 2, 58, 62, 63  
 randomized  
   block, 81  
   experiment, 2–5, 57, 62  
   recruitment, 108  
  
 rate ratio, 13, 20, 36, 54, 81, 165, 194,  
   208  
 recall bias, 25, 58, 164, 214  
 recurrent events, 162, 170  
 regression  
   adjustment, 4, 15, 16, 82, 83, 95, 96,  
     109, 115  
   calibration, 228, 238  
 relative risk, 13, 34, 193, 196  
 retrospective study, 2, 8, 57  
 reverse causation, 25  
 risk  
   difference, 102  
   ratio, 13, 17, 34, 35, 59, 193, 196  
   set, 20, 80, 162, 164, 165, 172, 180,  
     187, 199  
 Rubin's rules, 185, 189  
  
 sample size, choice of, 45  
 sampled risk set, 165, 166, 183, 184, 187,  
   201  
 sampling, 132  
   with replacement, 187  
 sampling bias, 89, 119  
 sampling model, 42, 43, 46, 47, 67, 76,  
   78, 79, 83, 85, 88, 100, 113, 114,  
   121  
 sandwich formula, 88, 116, 144, 182,  
   197, 201, 257, 261  
 scale of effort involved in planning  
   investigation, 44  
 score function, 252  
 secondary study base, 21  
 selection bias, 23, 30, 241  
 self-control by cases, 124  
 self-controlled case-series method, 125,  
   128, 131  
 sensitivity of procedure, 214, 221  
 shaking the data, 237  
 sibling controls, 112  
 significance test, 256  
 simulation extrapolation, 237  
 smoking, 26, 27, 34, 59  
   and lung cancer, 26, 61, 251  
 social class and discrimination by  
   prospective employers, 79  
 social science study in career choice, 27  
 sociology, 26, 79  
 specificity of procedure, 214, 221  
 state failure, 28  
 stochastic process model, 128

- stratified
  - analysis, 14
  - case-subcohort study, 204
  - sampling, 51, 85, 89, 109, 132, 136, 158
- stroke and risk factor studies, 112
- sufficient statistic, 37, 59, 68, 91
- sunlamp use and melanoma, 27, 41
- surrogate exposure, 171
- survivor function, 80
- synthesis of studies, 240
- synthetic retrospective study, 161
- systematic errors, 23, 226
  
- telephone interviews, use of, 213
- temporal aspects, 54, 80, 160
- three-stage
  - sample, 150
  - study, 157
- tied event times, 165
- time, incorporation into selection of controls, 18–20, 54–57, 80
- time-dependent exposure, 170
- traffic accident studies, 125
  
- transient exposure, 124, 127
- twin studies, 16, 154
- two-by-two table, 12, 17, 35, 59, 82, 122
- two-phase sample, 133
- two-stage design, 132, 134, 137, 157, 158
- two-to-one matching, 77
  
- uninformative prior, 259
- unmatched study, 10
  
- vaccination and self-control studies, 124, 128
- vector score, 253
- veterinary study, 28
  
- Weibull distribution, 182
- weighted
  - analysis, 158
  - least squares, 50, 61, 241
  - likelihood, 158
  - partial likelihood, 174
- weights, 200
  - calibrated, 207
  - estimated, 207