

Index

- α , *see* significance level
- β_2 , *see* kurtosis
- γ , *see* unnormalized skewness
- γ_1 , *see* skewness
- γ_2 , *see* excess
- ε , *see* efficiency
- μ . *see* Gaussian average value *or* signal strength
- ρ , *see* correlation coefficient
- σ , *see* standard deviation *or* Gaussian standard deviation
- τ , *see* lifetime
- Φ , *see* Gaussian cumulative distribution
- χ^2
 - distribution, 32
 - method, 114
 - binned case, 119
 - in multiple dimensions, 132
 - random variable, 32, 114, 120
 - Baker–Cousins, 120
 - Neyman’s, 119
 - Pearson’s, 119
- Ω , *see* sample space
- 3σ evidence, 207
- 5σ observation, 207

- activation function, 191
- adaptive boosting, 198
- AI, artificial intelligence, 195
- alternative hypothesis, 175
- Anderson–Darling test, 184
- Argus function, 43
- artificial intelligence, 195
- artificial neural network, 181, 190
 - deep learning, 192
- Asimov dataset, 231
- asymmetric errors, 99, 110
 - combination of, 123
- asymptotic formulae for test statistics, 231
- average value
 - continuous case, 27
 - discrete case, 12
 - in Bayesian inference, 69

- back propagation, neural network, 191
- background
 - dependence of Feldman–Cousins upper limits, 218, 220
 - determination from control regions, 129
 - fluctuation for significance level, 205
 - in convolution and unfolding, 160
 - modeling in extended likelihood, 107
 - modeling with Argus function, 43
 - rejection in hypothesis test, 176
 - treatment in iterative unfolding, 171
 - uncertainty in significance evaluation, 209
 - uncertainty in test statistic, 227, 236
- Baker–Cousins χ^2 , 120
- Bayes factor, 73
- Bayes’ theorem, 59
 - learning process, 67
- Bayesian
 - inference, 68
 - probability, 59, 64
 - visual derivation, 60
 - unfolding, 166

- BDT, *see* boosted decision trees
- Bernoulli
 - probability distribution, 17
 - random process, 16
 - random variable, 17
- Bertrand's paradox, 7
- best linear unbiased estimator, 133
 - conservative correlation assumption, 137
 - intrinsic information weight, 136
 - iterative application, 139
 - marginal information weight, 136
 - negative weights, 135, 137
 - relative importance, 136
- beta distribution, 83
- bias, 102
 - in maximum likelihood estimators, 113
- bifurcated Gaussian, 124
- bimodal distribution, 14
- bin migration, 158
- binned Poissonian fit, 120
- binning, 118
 - in convolution, 158
- binomial
 - coefficient, 18
 - interval, 147
 - probability distribution, 18, 147
 - Poissonian limit, 40
 - random process, 17
 - random variable, 18
- BLUE, *see* best linear unbiased estimator
- boosted decision trees, 181
 - adaptive boosting, 198
 - boosting, 198
 - cross entropy, 196
 - decision forest, 197
 - Gini index, 196
 - leaf, 196
 - node, 196
- boosting, boosted decision trees, 198
- Box–Muller transformation, 89
- Brazil plot, 225
- breakdown point, robust estimator, 103
- Breit–Wigner
 - non-relativistic distribution, 41
 - relativistic distribution, 42

- Cauchy distribution, 41
- central
 - interval, 70
 - limit theorem, 46
 - value, 99
- chaotic regime, 82
- classical probability, 4

- Clopper–Pearson binomial interval, 147
- CL_s method, 221
- CNN, *see* convolutional neural network
- coefficient of determination R^2 , 117
- combination
 - of measurements, 129
 - principle, 136, 140
- conditional
 - distribution, 53
 - probability, 9
- confidence
 - interval, 100, 109, 143
 - level, 100
- conservative
 - CL_s method, 221, 223
 - correlation assumption, BLUE method, 137
 - interval, 147
 - limit, 217
- consistency of an estimator, 102
- control
 - region, 129
 - sample, 130
- convergence in probability, 22
- ConvNet, *see* convolutional neural network
- convolution, 155
 - convolutional neural network, 194
 - Fourier transform, 156
 - convolutional neural network, 193
 - feature map, 194
 - local receptive fields, 194
- correlation coefficient, 14
- counting experiment, 208, 212, 216, 227
- Cousins–Highlands method, 227
- covariance, 14
 - matrix, 14
- coverage, 100
- Cramér–Rao bound, 102
- Cramér–von Mises test, 184
- credible interval, 70
- cross entropy, decision tree, 196
- Crystal Ball function, 44
- cumulative distribution, 28
- cut, 176

- data sample, 99
- decision
 - forest, 197
 - tree, 196
- deep learning, artificial neural network, 192
- degree of belief, 65
- dices, 4–6, 16, 21
- differential probability, 26
- discovery, 205, 207, 208

- distribution, *see* probability distribution
- dogma, extreme Bayesian prior, 66
- `drand48` function from C standard library, 84
- efficiency
 - hit-or-miss Monte Carlo, 90
 - of a detector, 10, 158
 - estimate, 104
 - of an estimator, 102
- elementary event, 4, 6, 9
- equiprobability, 4, 6, 25
- ergodicity, 94
- error
 - of a measurement
 - Bayesian approach, 70
 - frequentist approach, 99
 - of the first kind, 177
 - of the second kind, 177
 - propagation
 - Bayesian case, 79
 - frequentist case, 121
 - simple cases, 121
- estimate, 68, 97, 99, 100
- estimator, 100
 - efficiency, 102
 - maximum likelihood, 105
 - properties, 101
 - robust, 103
- Euler characteristic, 246
- event, 2
 - counting experiment, 187, 206
 - elementary, 4, 6, 9
 - in physics, 105
 - in statistics, 2
 - independent, 10
- evidence
 - 3σ significance level, 207
 - Bayes factor, 73
- excess, 15
- exclusion, 211
- expected value. *see* average value
- exponential distribution, 34
 - random number generator, 87
- extended likelihood function, 106, 186
- fast Fourier transform, 156
- feature map, convolutional neural network, 194
- feedforward multilayer perceptron, 190
- Feldman–Cousins unified intervals, 152
- FFT. *see* fast Fourier transform
- Fisher information, 75, 102, 136
- Fisher’s linear discriminant, 178
- flat (uniform) distribution, 6, 30
- flip-flopping, 150
- forest, boosted decision trees, 197
- Fourier transform of PDF convolution, 156
- frequentist
 - inference, 100
 - probability, 3, 22
- full width at half maximum, 31, 41
- fully asymmetric interval, 70
- FWHM, full width at half maximum, 31
- gamma function, 33, 71
- Gaussian
 - average value, μ , 31
 - bifurcated, 124
 - contours in two dimensions, 55
 - cumulative distribution, 31
 - distribution, 31
 - in more dimensions, 54
 - intervals, 32, 58
 - likelihood function, 108
 - random number generator, 89
 - central limit theorem, 88
 - standard deviation, σ , 31
- generator, *see* pseudorandom number generator
- Gini index, decision tree, 196
- global significance level, 242
- goodness of fit, 33, 118, 120
- `gsl_rng_rand` function from GSL library, 84
- Hastings ratio, 93
- histogram, 119
 - convolution, 158
 - in Asimov dataset, 231
 - PDF approximation, 182
- hit-or-miss Monte Carlo, 90
- homogeneous Markov chain, 93
- Hui’s triangle, 18, 23
- hybrid frequentist approach, 227
- hypothesis test, 175
- IID, independent identically distributed
 - random variables, 82
- IIW, intrinsic information weight, BLUE
 - method, 136
- importance sampling, 91
- improper prior distribution, 76
- incomplete Gamma function, 41

- independent
 - and identically distributed random variables, 82, 106
 - events, 10
 - random variables, 50
- inference, 97
 - Bayesian, 68
- intersubjective probability, 75
- intrinsic information weight, BLUE method, 136
- invariant prior, *see* Jeffreys' prior
- iterative unfolding, 166

- Jeffreys' prior, 75
- joint probability distribution, 49

- kernel function, *see* response function
- Kolmogorov distribution, 183
- Kolmogorov–Smirnov test, 182
- kurtosis, 15
 - coefficient, 15

- L'Ecuyer pseudorandom number generator, 84
- L*-curve, 165
- Lüscher pseudorandom number generator, 84
- Landau distribution, 46
- large numbers, law of, 21
- law
 - of large numbers, 21
 - of total probability, 11
- leaf, decision tree, 196
- learning
 - process in Bayesian probability, 67
 - rate parameter, artificial neural network, 191
- least squares method, 114
- lifetime, 35, 39
 - Bayesian inference, 76
 - Jeffreys prior, 77
 - maximum likelihood estimate, 112
 - measurement combination, 140
- likelihood
 - function, 67, 105
 - extended, 106, 186
 - Gaussian, 108
 - in Bayesian probability, 67
 - ratio
 - discriminant, 181
 - in search for new signals, 185, 209
 - projective discriminant, 182
 - test statistic in Neyman–Pearson lemma, 181
- linear regression, 115
- local
 - receptive fields, 194
 - significance level, 210, 242
- log normal distribution, 33
- logistic map, 82
- look elsewhere effect, 210, 242
 - in more dimensions, 246
- Lorentz distribution, 41
- loss function, 191
- lower limit, 70
- `rand48` function from C standard library, 84

- machine learning, 188
 - observation, 189
 - supervised, 188
 - unsupervised, 188
- marginal
 - distribution, 49
 - information weight, BLUE method, 136
- Markov chain, 93
 - homogeneous, 93
 - Monte Carlo, 69, 93
- maximum likelihood estimator, 105
 - bias, 113
 - properties, 112
 - method, 69, 105
 - uncertainty, 109
- MC, *see* Monte Carlo
- MCMC, Markov chain Monte Carlo, 93
- median, 14, 28, 103
- Mersenne-Twister pseudorandom number generator, 84
- Metropolis–Hastings
 - algorithm, 93
 - proposal distribution, 93
 - ratio, 95
- minimum
 - χ^2 method, *see* χ^2 method
 - variance bound, 102
- MINUIT, 106, 110
- misidentification probability, 176
- MIW, marginal information weight, BLUE method, 136
- mode, 14, 28
- modified frequentist approach, 221
- Monte Carlo method, 6, 46, 69, 81
 - hit-or-miss, 90
 - numerical integration, 92
 - sampling, 89
- multilayer perceptron, 190
- multimodal distribution, 14, 28

- multinomial distribution, 20
- multivariate analysis, 178
- MVA, multivariate analysis, 178

- negative weights, 135, 137
- nested hypotheses, *see* Wilks' theorem
- neural network, *see* artificial neural network
- Neyman
 - confidence belt
 - binomial case, 147
 - construction, 144
 - Feldman–Cousins, 152, 218
 - Gaussian case, 146
 - inversion, 146
 - confidence intervals, 215
- Neyman's χ^2 , 119
- Neyman–Pearson lemma, 181
- node, decision tree, 196
- normal
 - distribution, *see* Gaussian distribution
 - random variable, 31
- normalization condition, 9, 26
- nuisance parameter, 69, 98, 226, 227
- null hypothesis, 175

- observation in machine learning, 189
- observation, 5σ significance level, 207
- odds
 - posterior, 64, 73
 - prior, 73
- ordering rule, 144
- outlier, 103
- overcoverage, 100, 144, 146, 147, 217

- p -value, 118, 206
- parameter
 - estimate, 100
 - Bayesian, 68
 - nuisance, 226, 227
 - of interest, 69, 98
- Pascal's triangle, 23
- PDF, *see* probability distribution function
- Pearson's χ^2 , 119
- percentile, 28
- period of a pseudorandom number generator, 84
- POI, parameter of interest, 69, 98
- Poisson distribution, 35
 - Gaussian limit, 40
- Poissonian, *see* Poisson distribution
 - random variable, 35

- pooling, convolutional neural network, 194
- posterior
 - odds, 64, 73
 - probability, 60, 65, 67
- prior
 - odds, 73
 - probability, 60, 65
 - distribution, 67
 - distribution, improper, 76
 - distribution, uniform, 71, 74
 - subjective choice, 74
 - uninformative, 69, 75
- probability, 2
 - axiomatic definition, 8
 - Bayesian, 3, 4
 - classical, 4
 - density, 25
 - dice rolls, 5
 - distribution, 9, 25
 - χ^2 , 32
 - Bernoulli, 17
 - beta, 83
 - bimodal, 14
 - binomial, 18
 - Breit–Wigner, non-relativistic, 41
 - Breit–Wigner, relativistic, 42
 - cumulative, 28
 - exponential, 34
 - Gaussian, 31
 - Gaussian, in more dimensions, 54
 - joint, 49
 - Landau, 46
 - log normal, 33
 - Lorentz, 41
 - marginal, 49
 - multimodal, 14, 28
 - multinomial, 20
 - normal, 31
 - Poissonian, 35
 - standard normal, 31
 - uniform, 6, 30
 - distribution function, 26
 - in more dimensions, 49
 - frequentist, 3, 22
 - posterior, 60, 65, 67
 - prior (*see* prior probability)
 - theory, 2
- profile likelihood, 185, 228
- projective likelihood ratio discriminant, 182
- pseudorandom number, 81
 - generator, 82
 - drand48 function, 84
 - gsl_rng_rand function, 84
 - lrand48 function, 84

- exponential, 87
 - from cumulative inversion, 86
 - Gaussian, Box–Muller, 89
 - Gaussian, central limit theorem, 88
 - L’Ecuyer, 84
 - Lüscher, 84
 - Mersenne-Twistor, 84
 - period, 84
 - RANLUX, 84
 - seed, 84
 - uniform, 84
 - uniform on a sphere, 87
- purity, 64, 196

- quantile, 28

- R^2 , coefficient of determination, 117
- RANLUX pseudorandom number generator, 84
- random
 - number, *see* pseudorandom number generator, *see* pseudorandom number generator
 - process, 2
 - variable, 4
 - χ^2 , 32
 - Bernoulli, 17
 - binomial, 18
 - exponential, 34
 - Gaussian, 31
 - independent, 50
 - log normal, 33
 - normal, 31
 - Poissonian, 35
 - standard normal, 31
 - uncorrelated, 14
 - uniform, 6, 30
- Random forest, 197
- rate parameter, 35
- receiver operating characteristic, 177, 200
- rectified linear units, convolutional neural network, 195
- reference analysis, 76
- reference prior, 76
- regularization strength, 164
- regularized unfolding, 163
- relative importance, BLUE method, 136
- religious belief, Bayesian extreme probability, 66
- ReLU, rectified linear units, 195
- repeatable experiment, 2
- residuals, 114
- resonance, 41

- response
 - function, 155
 - Gaussian case, 157
 - matrix, 158
- RMS, *see* root mean square
- robust estimator, 103
- ROC curve, receiver operating characteristic, 177
- ROOT, 106, 173
- root mean square, 13
- ROOUNFOLD, 173

- sample space, 8
- seed, pseudorandom number generator, 84
- selection, 176
 - efficiency, 176
- shortest interval, 70
- sigmoid function, 191
- signal
 - exclusion, 211
 - region, 129
 - strength, 185, 186
- signal-injected expected limit, 225
- significance level, 177, 205, 207–210
- simultaneous fit, 130
- singular value decomposition, 171
- skewness, 14
 - unnormalized, 15, 126
- smearing, 156
- sources of systematic uncertainty, 100
- standard
 - deviation, 13
 - continuous case, 27
 - normal
 - distribution, 31
 - random variable, 31
- statistical uncertainty, 100
- subjective probability, *see* Bayesian probability
- supervised machine learning, 188
- SVD, *see* singular value decomposition 171
- symmetric interval, 70
- systematic uncertainty, 99, 226, 227
 - sources, 100

- Tartaglia’s triangle, 23
- test
 - sample, 189
 - statistic, 176
 - for discovery, 230
 - for Higgs boson search, 231
 - for positive signal strength, 230
 - for upper limits, 230

- Tikhonov regularization, 164
- total probability, law of, 11
- toy Monte Carlo, 120
- training, 188
 - sample, 180, 182, 188
- transformation of variables, 15, 29, 121
 - Bayesian posterior, 79
- trial factor, 243
- trimmed average, 103
- TUNFOLD, 173
- type-I error, 177
- type-II error, 177

- uncertainty, 68, 70, 97, 99
 - interval, 99
 - Bayesian, 70
 - frequentist, 100
 - with maximum likelihood method, 109
- uncorrelated random variables, 14
- undercoverage, 100
- underfluctuation, 211
- unfolding, 155
 - L curve, 165
 - Bayesian, 166
 - bin-to-bin correction factors, 163
 - in more dimensions, 173
 - iterative, 166
 - regularization strength, 164
 - regularized, 163
 - response matrix inversion, 160
 - singular value decomposition, 171
 - Tikhonov regularization, 164
- unified intervals, Feldman–Cousins, 152, 218
- uniform
 - distribution, 6, 30
 - random number generator, 84
- uninformative prior, 69, 75
- unknown parameter, 97
- unnormalized skewness, 15, 126
- unsupervised machine learning, 188
- upcrossing, 244
- upper limit, 70, 211, 215

- variance
 - continuous case, 27
 - discrete case, 13

- weighted average, 131
- Wilks' theorem, 120, 184
 - nested hypotheses, 184

- Z , *see* significance level