

Index

- S^2
- classical linear model 350, 354
 - MSE-Consistent 184
 - sampling distribution 165
 - unbiased 179
- \bar{X}
- definition 117
 - expectation and variance 117
 - independence of \bar{X} and S^2 165
 - MSE-Consistent 184
 - sampling distribution 165
 - standard error 182
 - unbiased 179
- σ -algebra 3, 9
- Borel σ -algebra 3
- a**
- asymptotic distribution 172
 - Central Limit Theorem 173
 - generalized likelihood ratio statistic 310
 - MLE 228
 - Score function 226, 316
 - asymptotic test
 - generalized likelihood ratio 310
 - Score test 316
 - Wald test 314
 - asymptotically unbiased estimator 181
- b**
- Bayesian estimation 247
 - Bayes estimator 251
 - Bayes posterior risk 251
 - Bayes risk 251
 - credible region 267, 268
 - hyperparameter 247
 - likelihood function 248
 - loss function 250
 - posterior distribution 247
 - posterior risk 251
 - prior distribution 247
 - risk function 250
 - setting 247
 - sufficient statistic 248
 - Bayesian estimator
 - absolute deviation loss 254
 - definition 251
 - squared error loss 252
 - beta distribution
 - mean 157
 - parameters 156
 - pdf 156
 - R commands 157
 - variance 157
 - bias 179
 - binomial
 - binomial expansion 36

- binomial distribution 129
 - mean 130
 - moment generating function 130
 - multinomial marginal 139
 - pdf 130
 - R commands 132
 - setting 130
 - variance 130
- Birthday Problem 41
- bootstrap method 277
 - algorithm 277
 - confidence interval 278
 - standard error 278
- C**
- Case Study
 - Coin Tosses 123
 - IDNAP experiment 375
 - Modeling Obsidian Rind Thicknesses 273
 - The Birthday Problem 41
 - The Central Limit Theorem 172
 - The Order Statistics 214
 - Titanic Survival 323
- CDF, *see* cumulative distribution function 47
- Central Limit Theorem 172, 173
 - for sums 175
 - proof 173
- chance experiment 1
- Chi-square distribution 150
 - S^2 165
 - degrees of freedom 150
 - mean 150
 - moment generating function 150
 - pdf 150
 - R commands 151
 - variance 150
- classical linear model 348
 - t -test 360
 - complete sufficient statistic 349
 - components 348
 - confidence interval 360
 - design matrix 348
 - estimability 352
 - estimable linear function 352
 - extra sums of squares test 360
 - least squares estimator 350
 - MLE 349
 - MLE properties 353
 - normal equations 349
 - residual deviance 357
 - sampling distribution S^2 356
 - sampling distribution SS 's 358
 - sampling distributions 355
 - simple linear regression 350
 - SSE 358
 - SSM 358
 - UMVUE's 354
- complete statistic 199
- composite hypothesis 281
- conditional distribution 57
 - expectation 106
 - pdf 58
 - variance 108
- conditional probability 19, 20
 - Bayes' Theorem 23
 - rules 21
- confidence interval 258, 259
 - classical linear model 360
 - confidence level 259
 - coverage probability 258
 - for μ 169
 - generalized linear model 342
 - logistic regression 369
 - no pivotal quantity 262
 - pivotal quantity 260
- confidence region 260
 - p %likelihood region 267
- consistent estimator 184
- convergence in distribution 172
 - definition 173
- correlation
 - correlation coefficient 103
 - independent random variables 105
 - properties 104

- uncorrelated 103
- counting rules 17, 34
 - allocation to cells 38
 - combination 36
 - Fundamental Principle of Counting 33
 - number of combinations 36
 - number of permutations 36
 - permutation 35
- covariance 99
 - covariance matrix 105
 - definition 99
 - independent random variables 101
 - properties 100, 101
- Cramér–Rao lower bound
 - d -dimensional θ 211
 - for $\tau(\theta)$ 206
 - for θ 206
 - UMVUE 206
- critical region 282
 - generalized likelihood ratio test 306
 - Karlin-Rubin Theorem 298
 - MP 288
 - sufficient statistic 290
 - UMP 296
 - UMPU 300
 - unbiased 284, 290
- cumulative distribution function 47
 - marginal 57
 - properties 49
 - random vector 54, 55
- d**
 - Delta method 186
 - design matrix 348
 - deviance 338
 - extra deviance test 345, 373
 - null deviance 344
 - residual deviance 344, 357
 - scaled deviance 338
 - standardized deviance residual 371
 - digamma function 231
- distribution function, *see* cumulative distribution function 47
- e**
 - estimation
 - Bayesian 247
 - maximum likelihood estimation 226
 - MLE 226
 - point estimation 177
 - estimator
 - accuracy 181, 182
 - asymptotic efficiency 210
 - asymptotic relative efficiency 185
 - biased 179
 - Delta method 186
 - efficiency 210
 - Gaussian Kernel Density estimator 273
 - least squares 350
 - MAD 251
 - Mean Squared Error 182
 - MSE-Consistent 184
 - precision 182
 - Rao-Blackwellization 195
 - relative efficiency 185
 - risk 251
 - standard error 182
 - UMVUE 204
 - unbiased 179
 - event 3
 - disjoint events 4
 - independent events 28
 - mutually exclusive events 4
 - expectation 77
 - \bar{X} 117
 - column random vector 118
 - conditional
 - properties 107
 - conditional expectation 109
 - conditional mean 106
 - expected values 77
 - independent random variables 97

expectation (*contd.*)
 Law of Iterated Expectation 109
 linear combination 115
 moments 80
 properties 81
 random vector 94, 96
 vector of means 105
 explanatory variable 324
 exponential distribution 150
 mean 150
 moment generating function 150
 pdf 150
 R commands 151
 variance 150
 exponential family 195, 196
 $E[t(X)]$ 198
 canonical form 198
 complete sufficient statistic 199
 curved family 197
 generalized linear model 332
 likelihood function 197
 monotone likelihood ratio 297
 one-parameter family 196
 regular family 196
 sufficient statistic 197
 UMPU critical region 300

f

F distribution
 degrees of freedom 169
 F ratio 359
 pdf 169
 R commands 169
 sums of Chi-square random
 variables 170
 finite mixture model 274,
 275
 Fisher's Information 206
 alternative formula 208
 definition 207
 for an iid sample 207
 matrix 211

Fundamental Principal of Counting
 33

g

gamma distribution
 n th moment 149
 mean 149
 moment generating function 149
 pdf 149
 R commands 149
 scale parameter 149
 shape parameter 149
 variance 149
 gamma function 149
 properties 149
 Gaussian Kernel Density estimator
 273
 generalized likelihood ratio 306
 asymptotic distribution 310
 generalized likelihood ratio test
 generalized linear model 343
 logistic regression 371–373
 simple hypotheses 307
 generalized linear model 331
 canonical link 336
 classical linear model 348
 components 334
 confidence interval 342
 deviance 338
 exponential family 332
 extra deviance test 345, 373
 goodness of fit test 343
 hat matrix 371
 likelihood function 336
 linear predictor 333
 link function 333
 log-likelihood function 336
 logistic regression 366
 logistic regression model 334
 mean and variance 332
 model utility test 344
 nested models 339, 344
 null deviance 344

- Poisson regression 376
 - probit regression 366
 - R command 338
 - residual deviance 344
 - saturated model 338
 - Score function 337
 - simple logistic regression model 368
 - Wald test 342
 - geometric distribution 136
 - CDF 136
 - mean 137
 - MGF 137
 - pdf 136
 - variance 137
- h**
- hat matrix 371
 - hypergeometric distribution
 - mean 133
 - pdf 133
 - R commands 133
 - setting 132
 - variance 133
 - hypothesis testing 281
 - alternative hypothesis 282
 - alternative parameter space 282
 - classical linear model 360
 - composite hypothesis 281
 - critical region 282
 - extra deviance test 345, 373
 - extra sums of squares test 360
 - generalized likelihood ratio test 306, 343
 - goodness of fit test 343
 - hypothesis test 282
 - Karlin-Rubin Theorem 298
 - large sample test 310, 314
 - model utility test 344
 - most powerful test 288, 290
 - MP critical region 288
 - Neyman-Pearson Lemma 288
 - nuisance parameter 306
 - null hypothesis 282
 - null parameter space 282
 - p-value 284
 - parametric testing 281
 - power 283
 - power function 283
 - Score test 316, 317
 - simple hypothesis 281
 - size 283
 - statistical hypothesis 281
 - sufficient statistic 290
 - Type I error 282
 - Type II error 282, 283
 - UMP critical region 296
 - UMP test 296
 - UMPU critical region 300
 - unbiased critical region 284
 - Wald test 314
- i**
- iid random variables 65
 - independence
 - covariance 101
 - expectation 97
 - independent events 28
 - moment generating function 98
 - mutually independent events 29
 - pairwise independent events 29
 - random variables 63
 - indicator function 47
 - inequality 84
 - Bonferroni's Inequality 13
 - Boole's Inequality 13
 - Cauchy-Schwarz 102
 - Chebyshev's 85, 86
 - Jensen's 86
 - Markov's 85
 - interval estimation 258
 - confidence interval 258
 - coverage probability 258
 - credible region 268
 - interval estimator 258

k

- Kolmogorov 1
 - Kolmogorov's Axioms of Probability 9
 - Kolmogorov's probability function 9

l

- Lévy's Continuity Theorem 173
- large sample test 314
 - extra deviance test 345
 - generalized likelihood ratio test 310
 - goodness of fit test 343
 - model utility test 344
 - Score test 317
 - Wald test 314
- Law of Likelihood 223
- likelihood function 190
 - definition 190
 - dependent on sufficient statistic 224
 - exponential family 197
 - generalized linear model 336
- likelihood ratio
 - asymptotic distribution 310
 - generalized likelihood ratio 306
 - MLE 234
- linear combination
 - \bar{X} 117
 - expectation 115
 - moment generating function 119
 - variance 115–117
- linear predictor 333
- link function 333
 - binomial 366
 - canonical link 334
 - complementary log-log 366
 - logit 366
 - logit function 334
 - Poisson regression 376
 - probit 366

- log-likelihood function
 - definition 190
 - generalized linear model 336
 - logistic regression 367
 - not differentiable 233
 - log-normal distribution
 - mean 155
 - pdf 155
 - R commands 155
 - transformation to normal 156
 - variance 155
 - logistic regression 334, 366
 - deviance 371
 - goodness of fit test 372
 - hypothesis testing 371
 - log-likelihood 367
 - MLE's 367
 - model utility test 372
 - Score function 367
 - simple linear logistic regression 371
 - logit function 334
- m**
- marginal pdf, *see* probability density function 55
 - maximum likelihood estimation
 - classical linear model 349, 353
 - Maximum Likelihood Principle 226
 - multi-parameter models 235
 - one-parameter models 228
 - Wald test 314
 - maximum likelihood estimator
 - asymptotic properties 228
 - definition 226
 - function of a sufficient statistic 227
 - invariance property 227
 - mean 78
 - Mean Squared Error 182, 183
 - consistency 184
 - MGF, *see* moment generation function 86

- moment generating function 86
 - expansion in moments 87
 - properties 88
 - sum 119
 - sum of independent random variables 98
 - unique to CDF 89
- moments
 - definition 80
- monotone likelihood ratio 296
 - exponential family 297
- most powerful test 288
 - generalized likelihood ratio test 307
 - Neyman-Pearson Lemma 288
- multinomial distribution
 - binomial marginals 139
 - pdf 139
 - R commands 139
 - setting 139
- multiple comparisons 378
 - Bonferroni procedure 13, 378
 - experimentwise error rate 378
 - per-comparisonwise error rate 378
- n**
 - negative binomial distribution
 - alternate version 138
 - mean 137
 - moment generating function 137
 - pdf 136
 - R commands 136
 - setting 135
 - variance 137
 - Neyman-Fisher Factorization Theorem 191
 - Neyman-Pearson Lemma 288
 - normal distribution
 - Z transformation 153
 - Chi-square transformation 154, 155
 - mean 153
 - moment generating function 153
 - normal random variable 120
 - pdf 152
 - R commands 152
 - standard normal 152
 - variance 153
 - nuisance parameter 306
- o**
 - order statistic 215
 - k th order statistic 215
 - CDF $X_{(1)}$ and $X_{(n)}$ 217
 - definition 215
 - pdf of $X_{(1)}$ 217
 - pdf of $X_{(n)}$ 217
 - sufficiency of $(x_{(n)})$ 216
 - sufficiency of $x_{(1)}$ 216
 - sufficiency of $x_{(1)}$ and $(x_{(n)})$ 216
- p**
 - p-value 285
 - parameter space 177
 - restricted 234
 - pdf, *see* probability density function 46
 - pivotal quantity 260
 - definition 260
 - Poisson distribution
 - mean 134
 - moment generating function 134
 - pdf 134
 - R commands 134
 - variance 134
 - posterior distribution 126
 - prevalence probability 23
 - prior distribution 126, 247
 - probability
 - equally likely outcomes 16
 - Law of Total Probability 12
 - conditional form 22
 - general form 22
 - Multiplication Law 21
 - probability axioms 9
 - probability function 9

- probability (*contd.*)
 - probability measure 9
 - probability rules 9, 10
 - probability space 9
- probability density function
 - continuous marginal 56
 - continuous random variable 46, 47
 - discrete marginal 55
 - discrete random variable 46, 47
 - marginal 55
 - random vector 53, 54
 - range-dependent 214
- Probability Integral Transformation 72
- probability model 129
 - Bernoulli 130
 - beta 156
 - binomial 129
 - Chi-square 150
 - continuous 147
 - discrete 129
 - exponential 150
 - F distribution 169
 - finite mixture model 274
 - gamma 149
 - geometric 136
 - hypergeometric 132
 - log-normal 155
 - multinomial 138
 - negative binomial 135
 - normal 152
 - Poisson 134
 - T and F 166
 - two-parameter exponential 214
 - uniform 147
- r**
- R command
 - log-log regression 368
 - mixtools 276
 - probit regression 367
- R commands
 - beta 157
 - binomial 132
 - Chi-square 151
 - exponential 151
 - F 169
 - gamma 149
 - generalized linear model 338
 - hypergeometric 133
 - log-normal 155
 - logistic regression 367
 - MLE 231, 238
 - negative binomial 136
 - normal 152
 - normalmixEM 276
 - Poisson 134
 - T 167
 - uniform 147
- random sample 68
- random variable 45
 - continuous random variable 46
 - correlation 103
 - covariance 99
 - dependence 64
 - discrete random variable 46
 - expectation 78
 - expected value 77
 - first moment 78
 - independence 63
 - moments 77
 - quantile 49
 - standardized 84
 - standardized random variable 84
 - sum of random variables 163
- random vector 53
 - bivariate 53
 - covariance matrix 105
 - expectation 94
 - joint CDF 54
 - joint pdf 53, 54
 - vector of means 105
- Rao-Blackwell Theorem 194
- Rao-Blackwellization 195

regularity conditions 205
 response variable 324

S

sample 178
 random sample 177, 178
 sample space 1
 sampling distribution 178
 S^2 356
 \bar{X} 165
 classical linear model 355, 358
 F 170
 T 167, 168
 Score function 224, 316
 asymptotic distribution 226, 316
 generalized linear model 337
 logistic regression 367
 mean 225
 Score test 316, 317
 variance 225
 sensitivity 23
 set operation
 complementation 3
 De Morgan's Laws 5
 Distributive Laws 6
 families of sets 5
 intersection 3
 partition 6, 7
 set difference 4
 union 3
 simple hypothesis 281
 specificity 23
 standard deviation 82
 definition 82
 standard error 182
 \bar{X} 182
 asymptotic 228
 bootstrap 277, 278
 Delta method 186
 statistic
 complete 199
 definition 178
 estimator 178

 sufficient 190
 sufficient statistic 190
 critical region 290
 classical linear model 349
 complete 199
 definition 190
 exponential family 197
 factorization criterion 191
 Lehmann-Scheffé method 194
 maximum likelihood estimation
 227
 minimal sufficient 193, 194
 Neyman-Fisher Factorization
 Theorem 191
 one-to-one transformation 193
 Principle of Sufficiency 190
 Rao-Blackwell Theorem 194
 sufficient statistics 190
 sums of random variables 163
 discrete probability model 164
 iid normals 165
 iid standard normals 165
 normal distributions 165

t

T distribution
 comparison with standard normal
 167
 degrees of freedom 166
 iid normal random variables 168
 pdf 166
 R commands 167
 relationship to F 169
 T transformation 167
 transformations
 bivariate random vector 73
 cdf method 70
 continuous random variable 70,
 72
 discrete random variable 69
 Jacobian 73
 Probability Integral Transformation
 72

Type I error 282
Type II error 282

u

UMVUE 204
 classical linear model 354
 Cramér–Rao lower bound 206
 definition 204
 Lehmann–Scheffé Theorem 204
unbiased estimator 179
 S^2 179
 \bar{X} 179
 asymptotically unbiased 181
 classical linear model 353
 sum of two unbiased estimators
 181
uniform distribution
 mean 148
 moment generating function 148

pdf 147

R commands 147

variance 148

uniformly most powerful test 296
 critical region 298

v

variance 81

\bar{X} 117

 alternative formulas 82

 asymptotic 228

 column random vector 118

 conditional 108

 properties 108

 conditional variance 109

 definition 82

 linear combination 115–117

 properties 83

 sum of random variables 98, 102

