

# Contents

<b>1</b>	<b>Probability Theory</b> .....	1
1.1	Why Probability Matters to a Physicist .....	1
1.2	The Concept of Probability .....	2
1.3	Repeatable and Non-Repeatable Cases .....	2
1.4	Different Approaches to Probability .....	3
1.5	Classical Probability .....	4
1.6	Generalization to the Continuum .....	6
1.6.1	The Bertrand's Paradox .....	7
1.7	Axiomatic Probability Definition .....	8
1.8	Probability Distributions .....	9
1.9	Conditional Probability .....	9
1.10	Independent Events .....	10
1.11	Law of Total Probability .....	11
1.12	Average, Variance and Covariance .....	12
1.13	Transformations of Variables .....	15
1.14	The Bernoulli Process .....	16
1.15	The Binomial Process .....	17
1.16	Multinomial Distribution .....	20
1.17	The Law of Large Numbers .....	21
1.18	Frequentist Definition of Probability .....	22
	References .....	23
<b>2</b>	<b>Probability Distribution Functions</b> .....	25
2.1	Introduction .....	25
2.2	Definition of Probability Distribution Function .....	25
2.3	Average and Variance in the Continuous Case .....	27
2.4	Mode, Median, Quantiles .....	28
2.5	Cumulative Distribution .....	28
2.6	Continuous Transformations of Variables .....	29
2.7	Uniform Distribution .....	30

2.8	Gaussian Distribution .....	31
2.9	$\chi^2$ Distribution .....	32
2.10	Log Normal Distribution .....	33
2.11	Exponential Distribution.....	34
2.12	Poisson Distribution .....	35
2.13	Other Distributions Useful in Physics .....	41
	2.13.1 Breit–Wigner Distribution .....	41
	2.13.2 Relativistic Breit–Wigner Distribution.....	42
	2.13.3 Argus Function.....	43
	2.13.4 Crystal Ball Function .....	44
	2.13.5 Landau Distribution.....	46
2.14	Central Limit Theorem .....	46
2.15	Probability Distribution Functions in More than One Dimension .....	49
	2.15.1 Marginal Distributions.....	49
	2.15.2 Independent Variables .....	50
	2.15.3 Conditional Distributions.....	53
2.16	Gaussian Distributions in Two or More Dimensions .....	54
	References.....	58
<b>3</b>	<b>Bayesian Approach to Probability .....</b>	<b>59</b>
3.1	Introduction .....	59
3.2	Bayes’ Theorem.....	59
3.3	Bayesian Probability Definition .....	64
3.4	Bayesian Probability and Likelihood Functions.....	67
	3.4.1 Repeated Use of Bayes’ Theorem and Learning Process .....	67
3.5	Bayesian Inference.....	68
	3.5.1 Parameters of Interest and Nuisance Parameters .....	69
	3.5.2 Credible Intervals .....	70
3.6	Bayes Factors .....	73
3.7	Subjectiveness and Prior Choice .....	74
3.8	Jeffreys’ Prior .....	75
3.9	Reference Priors .....	76
3.10	Improper Priors .....	76
3.11	Transformations of Variables and Error Propagation .....	79
	References.....	79
<b>4</b>	<b>Random Numbers and Monte Carlo Methods .....</b>	<b>81</b>
4.1	Pseudorandom Numbers.....	81
4.2	Pseudorandom Generators Properties.....	82
4.3	Uniform Random Number Generators .....	84
	4.3.1 Remapping Uniform Random Numbers .....	85
4.4	Discrete Random Number Generators .....	85

4.5	Nonuniform Random Number Generators.....	86
4.5.1	Nonuniform Distribution from Inversion of the Cumulative Distribution .....	86
4.5.2	Gaussian Generator Using the Central Limit Theorem .....	88
4.5.3	Gaussian Generator with the Box–Muller Method.....	89
4.6	Monte Carlo Sampling.....	89
4.6.1	Hit-or-Miss Monte Carlo .....	90
4.6.2	Importance Sampling .....	91
4.7	Numerical Integration with Monte Carlo Methods.....	92
4.8	Markov Chain Monte Carlo .....	93
	References.....	95
<b>5</b>	<b>Parameter Estimate .....</b>	<b>97</b>
5.1	Introduction .....	97
5.2	Inference.....	97
5.3	Parameters of Interest .....	98
5.4	Nuisance Parameters .....	98
5.5	Measurements and Their Uncertainties .....	99
5.5.1	Statistical and Systematic Uncertainties .....	99
5.6	Frequentist vs Bayesian Inference .....	100
5.7	Estimators .....	100
5.8	Properties of Estimators .....	101
5.8.1	Consistency .....	102
5.8.2	Bias .....	102
5.8.3	Minimum Variance Bound and Efficiency .....	102
5.8.4	Robust Estimators.....	103
5.9	Binomial Distribution for Efficiency Estimate .....	104
5.10	Maximum Likelihood Method .....	105
5.10.1	Likelihood Function .....	105
5.10.2	Extended Likelihood Function .....	106
5.10.3	Gaussian Likelihood Functions .....	108
5.11	Errors with the Maximum Likelihood Method .....	109
5.11.1	Second Derivatives Matrix .....	109
5.11.2	Likelihood Scan .....	110
5.11.3	Properties of Maximum Likelihood Estimators .....	112
5.12	Minimum $\chi^2$ and Least-Squares Methods .....	114
5.12.1	Linear Regression .....	115
5.12.2	Goodness of Fit and $p$ -Value .....	118
5.13	Binned Data Samples .....	118
5.13.1	Minimum $\chi^2$ Method for Binned Histograms .....	119
5.13.2	Binned Poissonian Fits .....	120

5.14	Error Propagation .....	121
5.14.1	Simple Cases of Error Propagation .....	121
5.15	Treatment of Asymmetric Errors .....	123
5.15.1	Asymmetric Error Combination with a Linear Model .....	124
	References .....	127
<b>6</b>	<b>Combining Measurements</b> .....	<b>129</b>
6.1	Introduction .....	129
6.2	Simultaneous Fits and Control Regions .....	129
6.3	Weighted Average .....	131
6.4	$\chi^2$ in $n$ Dimensions .....	132
6.5	The Best Linear Unbiased Estimator .....	133
6.5.1	Quantifying the Importance of Individual Measurements .....	135
6.5.2	Negative Weights .....	137
6.5.3	Iterative Application of the BLUE Method .....	139
	References .....	140
<b>7</b>	<b>Confidence Intervals</b> .....	<b>143</b>
7.1	Introduction .....	143
7.2	Neyman Confidence Intervals .....	143
7.2.1	Construction of the Confidence Belt .....	144
7.2.2	Inversion of the Confidence Belt .....	146
7.3	Binomial Intervals .....	147
7.4	The Flip-Flopping Problem .....	150
7.5	The Unified Feldman–Cousins Approach .....	152
	References .....	154
<b>8</b>	<b>Convolution and Unfolding</b> .....	<b>155</b>
8.1	Introduction .....	155
8.2	Convolution .....	155
8.2.1	Convolution and Fourier Transform .....	156
8.2.2	Discrete Convolution and Response Matrix .....	158
8.2.3	Efficiency and Background .....	158
8.3	Unfolding by Inversion of the Response Matrix .....	160
8.4	Bin-by-Bin Correction Factors .....	163
8.5	Regularized Unfolding .....	163
8.5.1	Tikhonov Regularization .....	164
8.6	Iterative Unfolding .....	166
8.6.1	Treatment of Background .....	171
8.7	Other Unfolding Methods .....	171
8.8	Software Implementations .....	173
8.9	Unfolding in More Dimensions .....	173
	References .....	173

- 9 Hypothesis Tests** ..... 175
  - 9.1 Introduction ..... 175
  - 9.2 Test Statistic ..... 175
  - 9.3 Type I and Type II Errors ..... 177
  - 9.4 Fisher’s Linear Discriminant ..... 178
  - 9.5 The Neyman–Pearson Lemma ..... 181
  - 9.6 Projective Likelihood Ratio Discriminant ..... 181
  - 9.7 Kolmogorov–Smirnov Test ..... 182
  - 9.8 Wilks’ Theorem ..... 184
  - 9.9 Likelihood Ratio in the Search for a New Signal ..... 185
  - 9.10 Multivariate Discrimination with Machine Learning ..... 188
    - 9.10.1 Overtraining ..... 189
  - 9.11 Artificial Neural Networks ..... 190
    - 9.11.1 Deep Learning ..... 192
    - 9.11.2 Convolutional Neural Networks ..... 193
  - 9.12 Boosted Decision Trees ..... 196
  - 9.13 Multivariate Analysis Implementations ..... 199
  - References ..... 203
- 10 Discoveries and Upper Limits** ..... 205
  - 10.1 Searches for New Phenomena: Discovery and Upper Limits ..... 205
  - 10.2 Claiming a Discovery ..... 206
    - 10.2.1  $p$ -Values ..... 206
    - 10.2.2 Significance Level ..... 207
    - 10.2.3 Significance and Discovery ..... 208
    - 10.2.4 Significance for Poissonian Counting Experiments ..... 208
    - 10.2.5 Significance with Likelihood Ratio ..... 209
    - 10.2.6 Significance Evaluation with Toy Monte Carlo ..... 210
  - 10.3 Excluding a Signal Hypothesis ..... 211
  - 10.4 Combined Measurements and Likelihood Ratio ..... 211
  - 10.5 Definitions of Upper Limit ..... 211
  - 10.6 Bayesian Approach ..... 212
    - 10.6.1 Bayesian Upper Limits for Poissonian Counting ..... 212
    - 10.6.2 Limitations of the Bayesian Approach ..... 215
  - 10.7 Frequentist Upper Limits ..... 215
    - 10.7.1 Frequentist Upper Limits for Counting Experiments ..... 216
    - 10.7.2 Frequentist Limits in Case of Discrete Variables ..... 217
    - 10.7.3 Feldman–Cousins Unified Approach ..... 218
  - 10.8 Modified Frequentist Approach: The  $CL_s$  Method ..... 221
  - 10.9 Presenting Upper Limits: The Brazil Plot ..... 225
  - 10.10 Nuisance Parameters and Systematic Uncertainties ..... 226
    - 10.10.1 Nuisance Parameters with the Bayesian Approach ..... 226
    - 10.10.2 Hybrid Treatment of Nuisance Parameters ..... 227
    - 10.10.3 Event Counting Uncertainties ..... 227

- 10.11 Upper Limits Using the Profile Likelihood ..... 228
- 10.12 Variations of the Profile-Likelihood Test Statistic ..... 229
  - 10.12.1 Test Statistic for Positive Signal Strength ..... 230
  - 10.12.2 Test Statistic for Discovery ..... 230
  - 10.12.3 Test Statistic for Upper Limits ..... 230
  - 10.12.4 Higgs Test Statistic ..... 231
  - 10.12.5 Asymptotic Approximations ..... 231
  - 10.12.6 Asimov Datasets ..... 231
- 10.13 The Look Elsewhere Effect..... 242
  - 10.13.1 Trial Factors ..... 243
  - 10.13.2 Look Elsewhere Effect in More Dimensions ..... 246
- References..... 248
  
- Index..... 251**



<http://www.springer.com/978-3-319-62839-4>

Statistical Methods for Data Analysis in Particle Physics

Lista, L.

2017, XVI, 257 p. 101 illus., 97 illus. in color., Softcover

ISBN: 978-3-319-62839-4