

Bayesian Reasoning in High Energy Physics  
- Principles and Applications -

Giulio D'Agostini  
Università di Roma "La Sapienza" and CERN

**CERN Yellow Report 99-03, July 1999**

## Abstract

Bayesian statistics is based on the intuitive idea that probability quantifies the degree of belief in the occurrence of an event. The choice of name is due to the key role played by Bayes' theorem, as a logical tool to update probability in the light of new pieces of information. This approach is very close to the intuitive reasoning of experienced physicists, and it allows all kinds of uncertainties to be handled in a consistent way. Many cases of evaluation of measurement uncertainty are considered in detail in this report, including uncertainty arising from systematic errors, upper/lower limits and unfolding. Approximate methods, very useful in routine applications, are provided and several standard methods are recovered for cases in which the (often hidden) assumptions on which they are based hold.

# Contents

<b>Introduction</b>	<b>1</b>
<b>I Subjective probability in physics? Scientific reasoning in conditions of uncertainty</b>	<b>3</b>
<b>1 Uncertainty in physics and the usual methods of handling it</b>	<b>5</b>
1.1 Uncertainty in physics . . . . .	5
1.2 True value, error and uncertainty . . . . .	6
1.3 Sources of measurement uncertainty . . . . .	7
1.4 Usual handling of measurement uncertainties . . . . .	8
1.5 Probability of observables versus probability of true values . . . . .	9
1.6 Probability of the causes . . . . .	10
1.7 Unsuitability of confidence intervals . . . . .	11
1.8 Misunderstandings caused by the standard paradigm of hypothesis tests . . . . .	13
1.9 Statistical significance versus probability of hypotheses . . . . .	16
<b>2 A probabilistic theory of measurement uncertainty</b>	<b>21</b>
2.1 Where to restart from? . . . . .	21
2.2 Concepts of probability . . . . .	22
2.3 Subjective probability . . . . .	24
2.4 Learning from observations: the ‘problem of induction’ . . . . .	26
2.5 Beyond Popper’s falsification scheme . . . . .	27
2.6 From the probability of the effects to the probability of the causes . . . . .	27
2.7 Bayes’ theorem for uncertain quantities: derivation from a physicist’s point of view . . . . .	29
2.8 Afraid of ‘prejudices’? Inevitability of principle and frequent practical irrelevance of the priors . . . . .	29
2.9 Recovering standard methods and short-cuts to Bayesian reasoning . . . . .	30
2.10 Evaluation of uncertainty: general scheme . . . . .	31
2.10.1 Direct measurement in the absence of systematic errors . . . . .	31
2.10.2 Indirect measurements . . . . .	34
2.10.3 Systematic errors . . . . .	34
2.10.4 Approximate methods . . . . .	37

<b>II Bayesian primer</b>	<b>39</b>
- slightly reviewed version of the 1995 DESY/Rome report -	
<b>3 Subjective probability and Bayes' theorem</b>	<b>41</b>
3.1 Original abstract of the primer . . . . .	41
3.2 Introduction to the primer . . . . .	41
3.3 Probability . . . . .	43
3.3.1 What is probability? . . . . .	43
3.3.2 Subjective definition of probability . . . . .	43
3.3.3 Rules of probability . . . . .	45
3.3.4 Subjective probability and objective description of the physical world . . . . .	47
3.4 Conditional probability and Bayes' theorem . . . . .	49
3.4.1 Dependence of the probability on the state of information . . . . .	49
3.4.2 Conditional probability . . . . .	49
3.4.3 Bayes' theorem . . . . .	51
3.4.4 Conventional use of Bayes' theorem . . . . .	53
3.4.5 Bayesian statistics: learning by experience . . . . .	54
3.5 Hypothesis test (discrete case) . . . . .	56
3.6 Choice of the initial probabilities (discrete case) . . . . .	57
3.6.1 General criteria . . . . .	57
3.6.2 Insufficient reason and maximum entropy . . . . .	59
<b>4 Distributions (a concise reminder)</b>	<b>63</b>
4.1 Random variables . . . . .	63
4.1.1 Discrete variables . . . . .	63
4.1.2 Continuous variables: probability density function . . . . .	65
4.1.3 Distribution of several random variables . . . . .	68
4.2 Central limit theorem . . . . .	71
4.2.1 Terms and role . . . . .	71
4.2.2 Distribution of a sample average . . . . .	72
4.2.3 Normal approximation of the binomial and of the Poisson distribution . . . . .	72
4.2.4 Normal distribution of measurement errors . . . . .	74
4.2.5 Caution . . . . .	74
<b>5 Bayesian inference applied to measurements</b>	<b>75</b>
5.1 Measurement errors and measurement uncertainty . . . . .	75
5.2 Statistical inference . . . . .	76
5.2.1 Bayesian inference . . . . .	76
5.2.2 Bayesian inference and maximum likelihood . . . . .	77
5.2.3 The dog, the hunter and the biased Bayesian estimators . . . . .	78
5.3 Choice of the initial probability density function . . . . .	79
5.3.1 Difference with respect to the discrete case . . . . .	79
5.3.2 Bertrand paradox and angels' sex . . . . .	79
5.4 Normally distributed observables . . . . .	81
5.4.1 Final distribution, prevision and credibility intervals of the true value . . . . .	81
5.4.2 Combination of several measurements . . . . .	82
5.4.3 Measurements close to the edge of the physical region . . . . .	83
5.5 Counting experiments . . . . .	85
5.5.1 Binomially distributed observables . . . . .	85

---

5.5.2	Poisson distributed quantities . . . . .	88
5.6	Uncertainty due to systematic errors of unknown size . . . . .	90
5.6.1	Example: uncertainty of the instrument scale offset . . . . .	90
5.6.2	Correction for known systematic errors . . . . .	91
5.6.3	Measuring two quantities with the same instrument having an uncertainty of the scale offset . . . . .	92
5.6.4	Indirect calibration . . . . .	94
5.6.5	Counting measurements in the presence of background . . . . .	94
<b>6</b>	<b>Bypassing Bayes' theorem for routine applications</b>	<b>97</b>
6.1	Approximate methods . . . . .	97
6.1.1	Linearization . . . . .	97
6.1.2	BIPM and ISO recommendations . . . . .	100
6.1.3	Evaluation of type B uncertainties . . . . .	101
6.1.4	Examples of type B uncertainties . . . . .	101
6.1.5	Caveat concerning the blind use of approximate methods . . . . .	103
6.2	Indirect measurements . . . . .	104
6.3	Covariance matrix of experimental results . . . . .	105
6.3.1	Building the covariance matrix of experimental data . . . . .	105
	Offset uncertainty . . . . .	106
	Normalization uncertainty . . . . .	107
	General case . . . . .	107
6.3.2	Use and misuse of the covariance matrix to fit correlated data . . . . .	108
	Best estimate of the true value from two correlated values. . . . .	108
	Offset uncertainty . . . . .	108
	Normalization uncertainty . . . . .	109
	Peelle's Pertinent Puzzle . . . . .	111
<b>7</b>	<b>Bayesian unfolding</b>	<b>113</b>
7.1	Problem and typical solutions . . . . .	113
7.2	Bayes' theorem stated in terms of causes and effects . . . . .	114
7.3	Unfolding an experimental distribution . . . . .	114
<b>III</b>	<b>Other comments, examples and applications</b>	<b>117</b>
<b>8</b>	<b>Appendix on probability and inference</b>	<b>119</b>
8.1	Unifying role of subjective approach . . . . .	119
8.2	Frequentists and combinatorial evaluation of probability . . . . .	120
8.3	Interpretation of conditional probability . . . . .	122
8.4	Are the beliefs in contradiction to the perceived objectivity of physics? . . . . .	123
8.5	Biased Bayesian estimators and Monte Carlo checks of Bayesian procedures . . . . .	125
8.6	Frequentistic coverage . . . . .	127
8.7	Bayesian networks . . . . .	128
8.8	Why do frequentistic hypothesis tests 'often work'? . . . . .	129
8.9	Frequentists and Bayesian 'sects' . . . . .	132
8.9.1	Bayesian versus frequentistic methods . . . . .	132
8.9.2	Orthodox teacher versus sharp student - a dialogue by Gabor . . . . .	133
8.9.3	Subjective or objective Bayesian theory? . . . . .	135

8.9.4	Bayes' theorem is not all . . . . .	137
8.10	Solution to some problems . . . . .	137
8.10.1	AIDS test . . . . .	137
8.10.2	Gold/silver ring problem . . . . .	138
<b>9</b>	<b>Further HEP applications</b>	<b>139</b>
9.1	Poisson model: dependence on priors, combination of results and systematic effects	139
9.1.1	Dependence on priors . . . . .	139
9.1.2	Combination of results from similar experiments . . . . .	140
9.1.3	Combination of results: general case . . . . .	141
9.1.4	Including systematic effects . . . . .	143
9.1.5	Is there a signal? . . . . .	145
9.1.6	Signal and background: a <i>Mathematica</i> example . . . . .	146
9.2	Unbiased results . . . . .	147
9.2.1	Uniform prior and fictitious quantities . . . . .	149
9.3	Constraining the mass of a hypothetical new particle: analysis strategy on a toy model . . . . .	150
9.3.1	The rules of the game . . . . .	150
9.3.2	Analysis of experiment <i>A</i> . . . . .	151
9.3.3	Naïve procedure . . . . .	151
9.3.4	Correct procedure . . . . .	153
9.3.5	Interpretation of the results . . . . .	154
9.3.6	Outside the sensitivity region . . . . .	155
9.3.7	Including other experiments . . . . .	157
<b>IV</b>	<b>Concluding matter</b>	<b>161</b>
<b>10</b>	<b>Conclusions</b>	<b>163</b>
10.1	About subjective probability and Bayesian inference . . . . .	163
10.2	Conservative or realistic uncertainty evaluation? . . . . .	164
10.3	Assessment of uncertainty is not a mathematical game . . . . .	165
	Acknowledgements . . . . .	166
	Bibliographic note . . . . .	166
	<b>Bibliography</b>	<b>169</b>

## Introduction

These notes are based on seminars and minicourses given in various places over the last four years. In particular, lectures I gave to graduate students in Rome and to summer students in DESY in the spring and summer of 1995 encouraged me to write the ‘Bayesian primer’, which still forms the core of this script. I took advantage of the academic training given at CERN at the end of May 1998 to add some material developed in the meantime.

Instead of completely rewriting the primer, producing a thicker report which would have been harder to read sequentially, I have divided the text into three parts.

- The first part is dedicated to a critical review of standard statistical methods and to a general overview of the proposed alternative. It contains references to the other two parts for details.
- The second part essentially reproduces the old primer, subdivided into chapters for easier reading and with some small corrections.
- Part three contains an appendix, covering remarks on the general aspects of probability, as well as other applications.

The advantage of this structure is that the reader can have an overall view of problems and proposed solutions and then decide if he wants to enter into details.

This structure inevitably leads to some repetition, which I have tried to keep to a minimum. In any case, *repetita juvant*, especially in this subject where the real difficulty is not understanding the formalism, but shaking off deep-rooted prejudices. This is also the reason why this report is somewhat verbose (I have to admit) and contains a plethora of footnotes, indicating that this topic requires a more extensive treatise.

A last comment concerns the title of the report. As discussed in the last lecture at CERN, a title which was closer to the spirit of the lectures would have been “Probabilistic reasoning . . . ”. In fact, I think the important thing is to have a theory of uncertainty in which “probability” has the same meaning for everybody: precisely that meaning which the human mind has developed naturally and which frequentists have tried to kill. Using the term “Bayesian” might seem somewhat reductive, as if the methods illustrated here would always require explicit use of Bayes’ theorem. However, in common usage ‘Bayesian’ is a synonym of ‘based on subjective probability’, and this is the reason why these methods are the most general to handle uncertainty. Therefore, I have left the title of the lectures, with the hope of attracting the attention of those who are curious about what ‘Bayesian’ might mean.

Email: [dagostini@roma1.infn.it](mailto:dagostini@roma1.infn.it)

URL: <http://www-zeus.roma1.infn.it/~agostini/>

