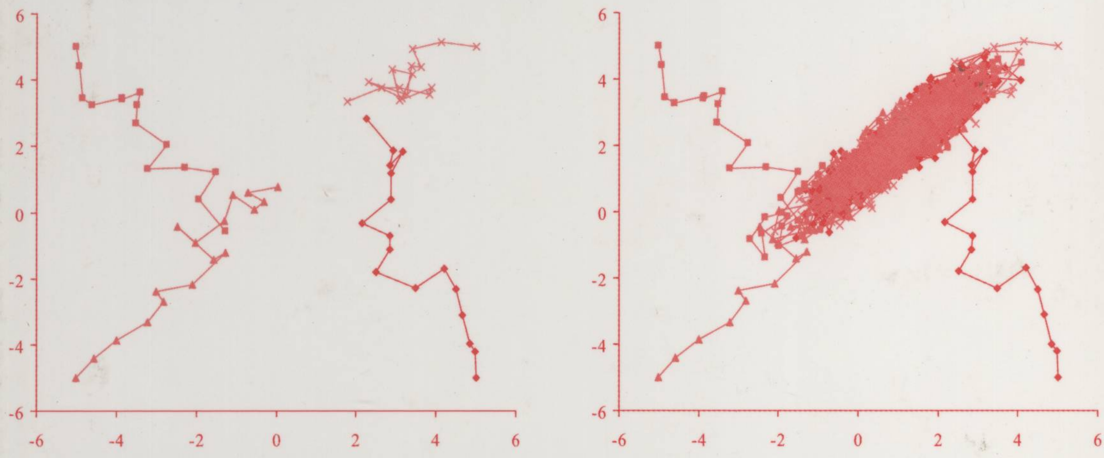
Bayesian Process Monitoring, Control and Optimization



edited by

Bianca M. Colosimo

Enrique del Castillo

Chapman & Hall/CRC
Taylor & Francis Group

Bayesian Process Monitoring, Control and Optimization

edited by
Bianca M. Colosimo
Enrique del Castillo

Chapman & Hall/CRC
Taylor & Francis Group
Boca Raton London New York

Chapman & Hall/CRC is an imprint of the
Taylor & Francis Group, an informa business

Chapman & Hall/CRC
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-10: 1-58488-544-0 (Hardcover)
International Standard Book Number-13: 978-1-58488-544-3 (Hardcover)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Bayesian Process Monitoring, Control and Optimization

Preface

During the last two decades, the use of Bayesian techniques has spread in different fields thanks to the advent of newly developed simulation-based approaches (e.g., Markov Chain Monte Carlo-MCMC) which have reduced computational barriers to the use of Bayesian inference in many applied fields. Problems that were considered intractable in the past are now routinely solved thanks to these approaches and to the ever increasing computer power. Furthermore, the long-lasting disjunctive about selecting "the best" approach, either Bayesian or frequentist, is now outdated in applied circles and substituted by a more pragmatic point of view which allows the analyst to select the most suitable approach depending on the problem at hand. During this time, the quality profession has also seen some applications and developments in Bayesian statistics, but this has occurred mainly in the more technical journals with little of this work reflected in actual practice.

While there are many excellent and recent books on applied Bayesian Statistics, most of them are related to biostatistics and Econometric applications. In engineering, the Bayesian approach has certainly been used by electrical and chemical engineers who have studied or applied Kalman filtering techniques (an area where there are several excellent textbooks too) for many years. However, the Bayesian paradigm is unfortunately not familiar to industrial engineers or to most applied or "industrial" statisticians, perhaps with the only exception of persons involved in reliability studies. Many industrial problems related to quality control and improvement require an in-depth use of statistical approaches, but in this era of "six-sigmas" and "black belts" the use of advanced statistical methods in practice is difficult to carry on. Despite the slight attention to Bayesian methodology in the industrial engineering and applied statistics technical journals, it can be observed, in general, a lack of attention to opportunities arising from the adoption of Bayesian approaches in actual industrial practice. We believe the gap between application and development can be bridged by having available more books on Bayesian statistics with a perspective on engineering applications.

It is with this goal that the project for the present book originated. The aim is to provide a state-of-the-art survey of applications of Bayesian statistics in three specific fields of industrial engineering and applied or industrial statistics, namely, process monitoring, process control (or adjustment), and process optimization. The book is intended as a reference for applied statisticians working in industry, process engineers and quality engineers working in manufacturing or persons in academia, mainly professors and graduate students in industrial and manufacturing engineering, applied statistics, and operations research departments. This is reflected in the diversity of the contributors to this book, who come from both academia and industry, and are located in the USA, Europe, and Asia.

The book is organized in four parts. Part I contains two introductory chapters. The first chapter provides an introduction to Bayesian statistics, emphasizing basic inferential problems, and outlines how these methods are applied in process monitoring, control, and adjustment. This chapter contains references and brief descriptions to all other chapters in the book, where appropriate. The second chapter presents a general overview of methods developed in the past few decades for computing Bayesian analysis via simulation (such as Markov Chain Monte Carlo and Monte Carlo simulation). The use of MCMC is illustrated with reference to a classic hierarchical model, the variance component model, using available software packages (WinBUGS and CODA, which runs under R). It is our hope that readers not familiar with Bayesian methodology will find in these two chapters a useful introduction and a guide to more advanced references.

Part II contains five chapters covering Bayesian approaches for process monitoring. The advantages of a Bayesian approach to process monitoring arise from the sequential nature of Bayes' theorem. As pointed out by some of the authors in this part of the book, a Bayesian approach allows a more flexible framework, in particular with respect to the usual assumption made in classical statistical process control (SPC) charts about known parameters. This part of the book deals with Bayesian methods for SPC and considers both univariate and multivariate process monitoring techniques. Application and development of full Bayesian approaches and empirical bayes methods are discussed.

The chapters in Part III present some Bayesian approaches which can be used for time series data analysis (for instance in case of missing data) and process control (also known as engineering process control). Here the use of the Kalman filter as an estimator of the state in a state-space formulation is exploited for prediction and control. This is perhaps the best known application of Bayesian techniques in engineering (outside of Industrial), although it is curious that Kalman himself did not develop his celebrated filter from a Bayesian point of view. Applications to radar detection and discrete part manufacturing are included.

Finally, Part IV focuses on Bayesian methods for process optimization. The three chapters included in this part of the book show how Bayesian methods can be usefully applied in experimental design and response surface methods (RSM). This section presents and illustrates the application of Bayesian regression to sequential optimization, the use of Bayesian techniques for the analysis of saturated designs, and the use of predictive distributions for optimization. The predictive approach to response surface optimization represents a major advance in RSM techniques, as it incorporates the uncertainty of the parameter estimates in the optimization process and has no frequentist counterpart.

We wish to thank all the contributing authors, with whom we share our interest in development of Bayesian statistics in industrial applications. A special thought goes to the late Carol Feltz (Northern Illinois University, USA) who showed praiseworthy spirit and strength in dedicating her last weeks to one chapter of this book.

Bianca M. Colosimo

Enrique del Castillo

Contributors

Frank B. Alt
University of Maryland
Robert H. Smith School
of Business
College Park, MD 20742
USA
falt@rhsmith.umd.edu

Marta.Y. Baba
School of Mathematical Sciences
Queen Mary, University of London
London E1 4NS
UK
myb@maths.qmul.ac.uk

Bianca M. Colosimo
Dipartimento di Meccanica
Politecnico di Milano
Piazza Leonardo da Vinci 32, 20133
Milano
Italy
biancamaria.colosimo@polimi.it

Enrique del Castillo
Department of Industrial &
Manufacturing Engineering
Penn State University
310 Leonhard Building
University Park, PA 16802
USA
exd13@psu.edu

Carol J. Feltz[1]
Division of Statistics
Northern Illinois University
DeKalb, IL 60115
USA

Steven G. Gilmour
School of Mathematical Sciences
Queen Mary, University of London
London E1 4NS
UK
s.g.gilmour@qmul.ac.uk

Spencer Graves
PDF Solutions, Inc.
333 West San Carlos, Suite 700
San José, CA 95126
USA
spencerg@pdf.com

Douglas M. Hawkins
Department of Statistics
University of Minnesota
313 Ford Hall
224 Church Street S.E.
Minneapolis, MN 55455
USA
doug@stat.umn.edu

Melinda Hock
Naval Research Laboratory
Washington, DC 20375
USA

[1]Sadly, Professor Carol Feltz passed away while completing chapter 4 of this book.

Carlos Moreno
Ultramax Corporation
110 Boggs Lane, Suite 325
Cincinnati, OH 45249
USA
carlos.moreno@ultramax.com

George Nenes
Aristotle University of Thessaloniki
Department of Mechanical Engineering
54124 Thessaloniki
Greece
gnenes@auth.gr

Rong Pan
Department of Industrial Engineering
Arizona State University
PO Box 875906
Tempe AZ 85287–5906
Rong.Pan@asu.edu

John J. Peterson
Statistical Sciences Department (UW281A)
GlaxoSmithKline Pharmaceuticals, R&D
709 Swedeland Road
King of Prussia, PA 19406–0939
USA
john.peterson@gsk.com

Jyh-Jen Horng Shiau
Institute of Statistics
National Chiao Tung University
Hsinchu
Taiwan
jyhjen@stat.nctu.edu.tw

Refik Soyer
School of Business
The George Washington University
Washington, DC 20052
USA
soyer@gwu.edu

George Tagaras
Aristotle University of Thessaloniki
Department of Mechanical Engineering
54124 Thessaloniki
Greece
tagaras@auth.gr

Panagiotis Tsiamyrtzis
School of Statistics
Athens University of Economics and Business
76 Patission Str, 10434
Athens
Greece
pt@aueb.gr

Contents

Part I

Introduction to Bayesian Inference

1

An Introduction to Bayesian Inference in Process Monitoring, Control and Optimization

Enrique del Castillo

Department of Industrial & Manufacturing Engineering, Pennsylvania State University

Bianca M. Colosimo

Dipartimento di Meccanica, Politecnico di Milano

CONTENTS

ABSTRACT We present in this first chapter a general overview of Bayesian inference. A brief account of the fundaments is given, after which we focus on the following problems: inference in normally distributed data (univariate and multivariate), Kalman filtering, and Bayesian linear regression. Applications are noted in process monitoring, control, and optimization. Whenever appropriate, we refer to all subsequent chapters in this book.

1.1 Introduction

This chapter presents a general description of Bayesian inference and assumes knowledge of undergraduate statistics. Appropriately, this description starts with telling who Bayes was and what his work was about. Unfortunately, not many biographical details are known about Bayes (for some of them, see Press [29]). Reverend Thomas Bayes, a Presbyterian minister, lived in England in the 18th century and wrote a manuscript on "inverse probability" related to making inferences about the proportion of a binomial distribution. This was published posthumously in 1763. In 1774, working independently, Laplace stated what is now known as Bayes' theorem in general form.

Bayesian inference combines prior beliefs about model parameters with evidence from data using Bayes' theorem. A subjective interpretation of probability exists in this approach, compared to the "frequentist" approach in which the probability of an event is the limit of a ratio of frequencies of events. The main criticisms of Bayesian analysis have been that it is not objective (a fact that has been debated for many years) and that the required computations are difficult. As it will be discussed in Chapter 2, the second criticism has been overcome to a large extent in the last 10 to 15 years due to advances in integration methods, particularly Markov Chain Monte Carlo (MCMC) methods. Interestingly, Fisher and other authors have speculated that Bayes' reluctance in publishing his manuscript was due to his own doubts about the principles behind it, in particular his interpretation of subjective probability and the specification of a prior distribution on the unknown parameter. However, Stliger [32] disagrees, indicating that Bayes was quite sure and direct about the use of subjective probability and the use of a prior distribution in his work, and that it was Bayes' difficulties with solving an integral (an incomplete beta function) that lead him not to publish. Thus, if Stigler is right, Bayes' himself would be very happy today about recent developments in the numerical solution of Bayesian inference problems.

In this chapter, we provide a succinct overview of Bayesian inference, mentioning applications in monitoring, control, and optimization of production processes. Whenever possible, we will make references to other contributed chapters in this volume.

1.2 Basics of Bayesian Inference

1.2.1 Notation

We first introduce some notation used in this chapter. Let θ denote unobserved quantities or population parameters of interest. This can denote a scalar or a vector of parameters as the context will make clear. In some sections, particularly on Kalman filtering, $\boldsymbol{\theta}$ will denote a vector of k parameters. Let y denote observable quantities (data), either a single observation or a vector of n observations $(y_1, \ldots, y_n)$ if the context does not require any other vector notation. If the context uses other vector notation (e.g., in multivariate data and regression analysis), then multivariate observations will be denoted by the column vector $\boldsymbol{y}$. If the data is time-ordered (as required in most process monitoring and control applications), we will use $Y^t = (y_t, y_{t-1}, y_{t-2}, \ldots)$ or $\mathbf{Y}^t = (\boldsymbol{y}_t, \boldsymbol{y}_{t-1}, \boldsymbol{y}_{t-2}, \ldots)$ to denote the data observed up to and including time instant t. The variable $\tilde{y}$ denotes a future observation of the same nature as y, and $\boldsymbol{X}$ denotes an $n \times k$ matrix of explanatory variables or covariates. $\boldsymbol{X}$ contains the "experimental design" in a regression analysis.

In this chapter, $p(\cdot)$ denotes a continuous density function and the notation $w|y$ denotes a conditional random variable w given y (the data). $P(\cdot)$ denotes the probability of some event defined over a sample space. Sometimes we will simply write "data" for all the data obtained from an experiment.

1.2.2 Goals of Bayesian Inference

The goal of Bayesian inference is to reach conclusions about — or perhaps, to make decisions based on — a parameter θ or future observation $\tilde{y}$ using probability statements *conditional* on the data y:

$$p(\theta|y) \rightarrow \text{posterior density of } \theta$$
$$p(\tilde{y}|y) \rightarrow \text{posterior predictive density of } y.$$

Bayesian inference considers all unknowns (parameters and future observations) as random variables. Classical (frequentist) statistical inference considers population parameters as fixed but data as random (due to sampling).

1.2.3 Bayes' Theorem for Events

Before looking at Bayes' theorem for densities, let us look first at the theorem for simple events. This should be familiar to any reader who has taken a basic probability course. For events A and B in some sample space $\mathcal{S}$, from the definition of conditional probability we have,

$$P(A|B) = \frac{P(A\cap B)}{P(B)}$$

and

$$P(B|A) = \frac{P(A\cap B)}{P(A)}.$$

These expressions are true if and only if

$$P(A\cap B) = P(A|B)P(B) = P(B|A)P(A)$$

from which

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}.$$

This formulation gives the essence of Bayes' theorem: if event B represents some additional information that becomes available, then $P(A|B)$ is the probability after this information becomes available, i.e., the *posterior* of A, and $P(A)$ is the probability before this information becomes available, i.e., the *prior* for A.

Suppose events A_i form a partition of $\mathcal{S}$, that is, $\bigcap_{\text{all}\,i} A_i = \mathcal{S}$; $A_i \cap A_j = \oslash$ for all $i, j, i \neq j$. Then the ("total") probability of B is given by

$$P(B) = \sum_{\text{all}\,j} P(B|A_j)P(A_j)$$

and therefore

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{\text{all}\,j} P(B|A_j)P(A_j)}. \tag{1.1}$$

Equation (1.1) is what Laplace [19] referred to as the problem of finding the "inverse probability" — given that the "effect" B is observed, find which of several potential "causes" was the true cause of the observed effect.

As clear from the above, nothing is incorrect in Bayes formula, as it is derived from the probability axioms. The debate concerns the interpretation of the probabilities involved. Classical statistics regards each probability in the