CLASSICAL AND BAYESIAN STATISTICS:
A SURVEY UPON THE DUALIST PRODUCTION

Paolo Rocchi[1] [§], Silvia Pandolfi[2], Lucia Rocchi[3]

[1]International Business Machines – IBM
53, Via Shangai, Roma, 00144, ITALY
e-mail: paolorocchi@it.ibm.com

[2]Department of Economics, Finance and Statistics
Via Pascoli, Perugia, 06123, ITALY
Universitá degli Studi di Perugia
e-mail: pandolfi@stat.unipg.it

[3]Department Economics, Appraisal and Food Science
Universitá degli Studi di Perugia
74, Via Borgo XX Giugno, Perugia, 06121, ITALY
e-mail: lucia.rocchi@unipg.it

Abstract: As the 20-th century advanced, the foundations of the so called *classical statistics* were laid and the *Bayesian statistics* experienced mounting popularity after the Second World War. Some statisticians firmly adhere to the classical or otherwise to the Bayesian methods, instead a third circle of dualist specialists accept both the views. Theoretical dualists attempt to merge the frequentist and the subjective conceptualizations together; in the meanwhile empirical dualists employ all the mathematical weapons in the living environment. The objective of this study is to analyze how the production of empirical dualists evolves over the decades, and how dualist authors see the fundamental questions which lie at the base of the frequentist and the subjective statistics. The present survey examines 452 electronic and paper books sampled from a University Library and Google Book Search. The results show the mounting popularity of the dualist movement and the substantial indifference of dualist authors to profound philosophical questions.

AMS Subject Classification: 62A03

[§]Correspondence author

1. Introduction

In the mind of most statisticians there are at least two mutually exclusive approaches to data analysis. The classical theory consisting of confidence intervals and hypothesis tests is used since the first half of the twentieth century and takes up the vast majority of the typical statistics tests. On the other hand Bayesian statistics attracts an ever increasing group of professionals. The classical and Bayesian methods are based on the frequentist and subjective interpretations of the probability. There is no universal agreement as to the precise definition of probability and the supporters of those schools fire fierce debates.

Aside radical authors we see a significant circle of statisticians who openly maintain that both the methods should be used. *Dualists* take this way as a good agreement between the opponents. They even consider the dualist approach as the loophole or the byproduct of the major intellectual approaches. Pragmatic writers produce journal articles and books in statistics to bring evidences of the advantages which a professional can obtain from the usage of different statistical methods.

Aside this pragmatic interpretation eminent theorists and philosophers such as Russel [14], Carnap [5], Popper [15], and more recently Gillies [9] develop accurate analysis and aim at demonstrating how multiple views on probability appear reasonable.

In conclusion we classify the dualist authors into two groups: a first circle includes *empirical dualist* authors and the second includes *theoretical dualist.*

Dualism has considerably developed in recent years and continues to increase in popularity, thus we aimed to conduct an investigation upon the movement of empirical dualists in statistics.

2. Research Design and Research Problems

During the course of the past decades some disputes amongst statisticians have been conciliated on the basis of practical considerations. As an example we quote the quarrel between Fisher and Pearson. In the 40s and 50s some authors rose above the problem and merged Fisher's theory of significance with the

Neyman-Pearson theory of testing, and presented this hybrid as statistics 'de facto' [17].

The dispute amongst Fisher, Neyman and Pearson was grounded upon shared theoretical statements, instead the empirical dualists have no reasonable basis. The astonishing sides of current empirical dualist movement are the following ones:

a) Frequentist and Bayesian radical claim how their approaches are absolutely incompatible.

b) Dualists have not yet devised a shared theoretical solution to conciliate the frequentist and the subjective interpretations, anyway empirical suggestions from the dualists seem to find mounting success in the professional practice.

Empirical dualists adopt methods that are clashing from the theoretical stance but their suggestions seem to work in the practice. This evident inconsistency between theory and practice drew our attention.

Whereas papers and articles sometimes develop conclusions that lie far away from professional practice, the books purchased and used by practitioners seem closer to the above mentioned contradiction theory/practice which working statisticians should meet. We are inclined to believe that the textbooks and the manuals mirror the factual impact of dualism in the living environment, thus we determined to investigate the production of pragmatic dualists.

The aim of this paper is to improve our knowledge on empirical dualism. In particular we mean to find out the dimension of the dualist movement and how empirical dualists tackle methodological problems despite the evident contradictions emerging in the points a) and b).

The major research questions are:

1. When did the dualist literature emerge and how did it change over the decades?

2. What are the most significant traits of dualist books?

3. How do dualist authors treat the fundamental questions which lie at the base of the frequentist and the subjective statistics?

3. Related Researches

We have not found researches on the dualist production so far, but two veins of research have some connections with the present inquiry.

1. Papers and journal articles authored by pragmatic dualists dissect the statistical problems which a practitioner can tackle with the classical approach or otherwise with the Bayesian approach. The writers encourage the use of multiple methods and discuss the most appropriate tool depending on the context of the problem. Among the earliest works we quote Box [4] who holds that the scientific method requires not one, but two kinds of inference. Box concludes that once this dualism is understood the statistical advances made in Bayesian methods, data analysis, robust and shrinkage estimators can be seen as a cohesive whole. Amongst the most recent papers we cite the joint contribution by Bayarri and Berger [3] who claim the classical and Bayesian approaches have a great deal to contribute to statistical practice. Each approach is actually essential for full development of the other approach. We briefly remind as dualist authors: G. Schafer, D.R. Cox, C. Chatfield, I. Levi, D. Costantini, L.E. Ballentine.

2. The popularity of the classical and Bayesian methods has been surveyed in various fields. Usually scientists wonder about the validity of each statistical approach in a certain area and survey upon the acceptance of the methods in use. Perhaps Hokanson is the most prolific researcher in this kind of inquiries; he surveyed the use of statistical techniques in oncology [12], in otolaryngology [13] and in psychiatry [11]. We remind the survey by Hakko and others over a sample of nearly thirty years [10] and the study by Emerson and Colditz on a local context [8]. We quote surveys on jurisprudence [16] and library science [7]. Wallace conducts a comparative analysis and finds that journals in library and information science produced more articles making no use of statistics than did journals in other subject areas [18].

4. Sources of Data

We meant to examine the books published in a significant arch of years. The Bayesian movement came to light after the Second World War and we fixed the sample should range form 1941 up to present: 2008. Because we were interested in the analysis of pragmatic dualism and its pervasiveness, the sample should include books that have wide commercial distribution. For such a reason we took a census of books presented in Google Book Search that fits this criterion. Google Book Search is a collection of electronic books delivered worldwide by major and minor publishers. This collection is served by Google search engine available in the web at http:// books.google.com and http://books.google.it. We made a census with incomplete list of all the statistical books present in the

Google collection, using the key words: 'introduction statistics', 'fundamentals probability', 'Bayesian statistics'. To avoid distortion made by Google and its criteria, which in some case reduce the access to books, we repeated our queries over almost one year to get stable results. We obtained 1,210 links to electronic books written in English.

Google Book Search offered statistical works created between 1971 and 2008. This representative sample did not cover the first steps of dualism movement. For such a reason, a future development of the present work is to considerer also traditional libraries, that have also books of former periods, but the research in traditional libraries, fundamental for the analysis of the dualist movement, is very complicated. In the present study, we considered also a traditional library such as an integration of the Google's books. We worked with the library of 'Scienze Giuridiche, Economiche ed Estimative' at the University of Perugia (Italy), a somewhat modern institution, that contains exclusively books on paper. The oldest books date back to the early nineteen-twenties. We used the catalogue of the library and examined all the entries classified under catalogue's section: 'Statistics'. The overall size of the section comes to 220 entries including books written in English and in Italian as well. The books from Perugia are not representative of the whole population of statistical books present, for example, in all the Italian university libraries but they are very useful to complete the analysis based only on Google. We consider books from Perugia as a test to verify the utility of a future research including also traditional library. Lastly we verified that the mounting trend of the overall sample mirrors the increasing production of books in the world during the same period [1].

The present survey aims at investigating the dualist authors, thus it was necessary to exclude the items that are not pertinent to this research from the overall population of 1,430 books. We adopted the following rules for exclusion of:

1. The books that do not treat statistical methods or any statistical topics e.g. Google Book Search exhibited texts on chemistry, manuals on the use of software packages and books on environmental sciences.

2. Statistical surveys, annual reports etc. which do not deal with fundamentals of statistics.

3. Proceedings, commemorative collections of papers and books whose chapters are written by different authors. This production is to be associated with research articles.

4. Redundant books which are the multiple editions of a title and the

variants of a title. What is a variant? Sometimes an author writes a top-seller work and soon after writes a variant of this work. We define B as a 'variant' of A whenever B describes the same topics of A; B replicates A in a way and is redundant. For example an author wrote: "An Introduction to Biostatistics: A Manual for Students in Health Science" and two years later published: "Introduction to Biostatistics and Research Methods" that basically deal with the same topics. Multiple editions and variants distort the present research since they inflate the impact of an author. Note that we do not consider redundant two books that differ in a way.

5. Books dealing with a topic that strictly pertains to a single school and a priori do not conform to the present survey on dualism. For example "Hierarchical Bayesian Optimization Algorithm" has been excluded. Instead we kept books which treat Bayesianism at large.

Using the above mentioned rules of exclusion we obtained a sample of 452 books:

a) 293 electronic books (65%) from Google Book Search.

b) 159 printed books (35%) from the library of the Perugia University [23 books written in Italian]

The rules of exclusion from 1 to 5 left out the vast majority of titles offered by Google (= 76%). Instead the second source turned out to be much more regular and involved basically rules 2 and 3, and 72% of the offered books resulted suitable for the present work. Books b) spans from the year 1941 to 2005, and a) ranges form 1971 up to 2008; the whole sample covers 68 years. Section b) shows more constant distribution over the years than the electronic books (see Figure 1). The library offers dated books on statistics and Google offers prevalently modern books thus these sources mutually compensate each other.

5. First Question Point

When did the dualist literature emerge and how did it change over the decades?

We analyzed the classes of the sampled books to reply to this question. We examined the historical trends and finally we saw the cultural inclination of dualist authors.

Classes

We defined the sampled books according to the following rules:

a) A book is classical if it exclusively deals with classical statistics.

b) A book is Bayesian if it exclusively deals with Bayesian statistics.

c) A book is dualist when:

— The author(s) treats classical statistical methods and Bayesian methods as well, no matter the topics appear unbalanced. We assume a dualist author can devote a few pages or even a few lines to one school and the remaining space to the other school. In principle the classical and the Bayesian methods regard the subject topics treated by the author in different ways.

— The author shows a positive and constructive description of both the statistical methods. The present feature is essential for dualism. We do not assume as dualist those who describe a statistical approach and quote the other approach exclusively to denigrate it or to extol their own approach. For example a large part of Bayesian authors comment on the frequentist approach but they have the clear and manifest purpose to show the advantages of Bayesianism. Thus we do not classify as dualist the Bayesians who describe the classical statistical method in order to put off this method. A dualist can even follow a methodology and reserve a small part of the available space to the other methodology but quotes this second methodology in a constructive manner.

We find the following results in accordance to the rules from a) to c):

1. 233 frequentist books,

2. 40 Bayesian books,

3. 179 dualist books.

Dualist books range nearly 40% versus 60% of remaining monothematic works, and this gross value highlights the significance of the dualist movement. Dualism production proves to be wide spreading.

We verified the degree of radicalism of non-dualists. We have seen 114 (49%) books authored by frequentists who do not quote the name of the Reverend Thomas Bayes and present his theorem as 'theorem on conditional statistics'. In addition we have found 6 books on classical statistics (0.02%) whose authors openly declare their position apart or against the Bayesian school. In conclusion classical statisticians show ignorance of Bayesianism in the early period of publication when Bayesian methods were not yet popularized and reveal certain indifference toward Bayesianism in recent years.

We have found 22 Bayesian books (56%) that quote classical statistics solely to disparage the conventional methods or anyway to underline the advantages of the Bayesianism; 33 books (84%) exhibit the terms 'Bayesian' or 'Bayes' in the title to underline the affiliation of the work. A small group of dualists - 20 writers - has Bayesian origin as we shall detail later. In conclusion Bayesians prevalently attempt to sustain their way to statistics and are not inclined toward dualism.

Main Historical Trends

The normalized production of books per year (Figure 2) shows how the advance of dualism may be subdivided into three periods.

The first period covers 26 years, from 1941 to 1966, and includes 19 frequentist books, 1 Bayesian book and there is no clue of dualism. According to Barlow [2] and others the birth of Bayesianism dates back to the De Finetti paper "La Prévision: Ses Lois Logiques, Ses Sources Subjectives" written in 1937. The Bayesian methods entered in use after the Second World War [6] and we find only one Bayesian book in this period while the frequentist school dominates the statistical contest.

The second period covers 14 years, from 1967 to 1980, and includes 26 frequentist books, 2 Bayesian books and 6 dualist books. The first signs of the dualism movement emerge while the traditional methods still dominate the period and Bayesianism emerges as an autonomous way. This may be defined as the transition period.

The third period covers 28 years, from 1981 to 2008, and includes 188 frequentist books, 37 Bayesian books and 173 frequentist books. The classical and Bayesian statistics are constantly present over the period while the popularity of the empirical movement appears noticeable. It seems reasonable to talk about the explosion of the dualist movement.

It's not possible to construct a synthetic index for each cultural movement but the annual densities for each period seem reasonable parameters to use. The annual density is the ratio between number of books and number of years in every period (see Table 1 and Table 2).

Annual densities have been used to construct a histogram when the class width is different in each period. The area of each bar is equal to the number of books (Figure 3). The percentage variation of dualist based on geometric average is 5.27% in the third period. The percentage increase of frequen-

PERIODS	Number of Books Annual			
	Dualist	Bayesian	Frequentist	TOTAL
1941-1966	0	1	19	20
1967-1980	6	2	26	34
1981-2008	173	37	188	398
	179	40	233	452

Table 1

PERIODS	Densities of Books		
	Dualist	Bayesian	Frequentist
1941-1966	0	0.038462	0.730769
1967-1980	0.428571	0.142857	1.857143
1981-2008	6.178571	1.321429	6.714286

Table 2

tist production between the third and the second periods reaches 6.23%, the Bayesians reach 1.750%, while the dualist production registers the astonishing value 2.783%.

Authoring

645 authors wrote the sampled books. We have analyzed how much authors cooperate to bring forth a single work but we have not found significant differences amongst the schools. The authors who wrote two or more books have been counted two or more times in this calculation.

	Dualist	Bayesian	Frequentist	
N. of Authors	272	69	337	Total = 678
Average N. of Authors/Book	1.52	1.73	1.45	Mean = 1.50

Table 3

Multi-authored statistical books are more popular in recent years in accor-

Period	First	Second	Third Average
N. of Authors/Book	1.10	1.24	1.54

Table 4

dance to the general trend in scientific publishing (see Table 4).

We have inquired the cultural extraction of authors and we have seen how the vast majority of writers refer to the traditional school on statistics and 93% of dualists have a classical extraction. Solely 20 dualist writers (7%) out of 281 have a Bayesian origin. In the sample we found 32 writers who authored two or more books. Twelve writers wrote a traditional book on statistics and later on have created a dualist text. A couple of Bayesian authors have written a Bayesian work and a dualist work as well. In total 14 writers out of 32 ($\sim$44%) were converted to dualism. We have not registered any other form of migration within the sample. In conclusion, dualism is seen as a cultural advance by statisticians who leave the monothematic methods toward the pluralist methods, while we do not see significant movements in the opposite way.

6. Second Question Point

What are the most significant traits of dualist books?

We have examined the topics developed by dualist authors, and the involvement of authors in the dualist culture to answer this question

Genres of Contents

The dualist books deal with a broad variety of topics. We have catalogued four genres according to the subjects treated in the books and to the structure of the books.

Class $\boldsymbol{\alpha}$ includes 43 textbooks and reference books (24% of dualist books) in probability and statistics used in schools and professional practice.

Class $\boldsymbol{\beta}$ has 56 books ($\sim$31%) similar to the previous but each work converges toward a special area. Disciplines involved with statistics rank the following range: business & economics (33.9%), medicine and psychology (25.0%), engineering & science (14.3%), biology (7.1%), environmental sciences (7.1%),

forensic sciences (5.4%), astronomy (3.6%), and sociology (3.6%).

Class γ contains 69 books ($\sim 38\%$) which examine advanced topics in statistics e.g. circular statistics, asymptotic statistics, statistical inference, bootstrap, or go deep into subtle arguments e.g. paradoxes and special cases in statistics.

Class δ includes 8 books ($\sim 5\%$) that have the structure of dictionaries, namely they do not have chapters instead they have entries; e.g. Statistics, Bayesian statistics, Bootstrap etc.

Class ε includes 3 books ($\sim 1\%$) that cannot be associated with the previous classes since they treat statistics in a lateral manner. In particular the purpose of {1} is to prepare the reader to AP Statistics Examination; {2} serves to teach and to set up lessons on statistics; {3} discusses more than 130 statistical tests (parametric; non-parametric etc.)

Involvement in Dualism

Dualist authors show a different awareness to the problems underlying multiple approaches to statistics. We have defined the degrees of participation to the dualist movement on the basis of the following procedure that includes two steps:

Step 1

1.1. We screen the books that devote a paragraph (or a number of pages equivalent to a paragraph) to discussing the crucial sides of the statistics foundations.

1.2. The selected books belong to Group I; for he remaining books see Step 2

Step 2

2.1. If the book basically follows one statistical school and dedicates more than 2 pages to the other approach, then the book belongs to Group II.

2.2. The remaining books belong to Group III.

The grouped books exhibit the ensuing characteristics:

Group I - 21 books (=12%) (11.4% of authors) display thorough involvement in dualism. The authors go into depth regarding the frequentist and the Bayesian methods and compare them. They analyze how the methods interrelate, what assumptions they are based on, and the practical implications of such distinctions. In particular the accurate discussion concerning the two schools covers one or more paragraphs in 11 books (52%), it covers a chapter (or sec-

tion) in 4 books (19%), 2 books develop a critical analysis in the introduction of the book, 3 books include the discussion in the conclusion of the book; 3 books make comments on scattered pages (Figure 4). The books prevalently fall into genres and (71.4%).

Group II - 116 works (=65%) (64,0% of authors) show a special style in the sense that writers focus on some special topics of the Bayesian statistics (e.g. prior/posterior probabilities, decision theory, estimators) and of the classical statistics (e.g. hypothesis testing, student-test) but do not develop a critical discussion of the statistical principles underling the two schools. Writers illustrate the frequentist and subjective ways but miss any debate. The following facts influence the author's behavior:

— Ninety per cent of the books are β, γ or δ namely pursue special scopes and deal with particular topics such as medicine and management science, or advanced topics in statistics. Authors are not involved in the complete account of statistics rather than they view statistics from a special perspective.

— A minority follows a special didactical style oriented to vulgarize the statistical topics. They arrange their writing according to a particular viewpoint and avoid going into non-technical questions.

In conclusion the books of Group II aim at highlighting the advantages offered by each mathematical tool beyond philosophical quarrels. This side of Group II will be analyzed in the next question point.

Group III - We have found 42 books (=23%) (24,6% of authors) that express the feeblest dualist position. They normally develop the frequentist mode (95%) and add a few lines or a few pages to mention the other school of thought. e.g. {25} spends half a page on subjectivism. e.g. {26} illustrates the Bayesian inference in a page. Basically authors of the present group focus on a statistical way and bunch the other solutions in a corner. They do not translate the dualist view into explicit directions because of the unbalanced illustration of the methods.

7. Third Question Point

How do dualist authors treat the fundamental questions lying at the base of the frequentist and subjective statistics?

We have gone deep into the contents of the dualist works I, II and III.

Group I

Authors prove to be conscious of the complex problems lying at the base of the multiple views on statistics and extensively debate the strong and weak points of each school. We have recognized two major groups:

I.1 Deep thinkers which are subdivided into profound commentators (7 authors), dualists as philosophers (4 authors), and fast commentators (11 authors).

I.2 Special writers subdivided into moderate Bayesians (3 authors), and original creators (4 authors).

Profound commentators

These authors are fully conscious of the dualism mode and even encourage the progress along this third way to statistics such as Chris Chatfield who offers an exhaustive view of both the fields and adds:

"This ecumenical or eclectic approach to statistics is one that I hope and expect to mature over the coming years (p. 92)".

Richard L. Smith, and G.A. Young openly declare that they present the concepts and results underlying the Bayesian, frequentist, and Fisherian approaches to statistical inference, with particular emphasis on the contrasts between them. They treat basic mathematical theory as well as more advanced topics as Bayesian computation, higher-order likelihood theory, predictive inference, bootstrap methods, and conditional inference.

Vic Barnett says:

"The object of this book is to examine the fundamental nature of statistical theory and practice by a comparative study of a different philosophical, conceptual and attitudinal (sometimes personal) 'approaches' to the subject".

And he develops methical parallel accounts of both the schools. Barnett is a well known dualist in the statistician community, cited in various papers.

B. H. Wonnacott, and Ronald J. Wonnacott authored a number of good books: 'Introduction to statistics', 'Student Workbook' etc. which we excluded from the sample due to the principle of redundancy. The present book is a clear treatise that includes several significant topics: robust estimation, non-linear and multiple regressions, Bayesian inference. This pair of writers should be included amongst the pioneers of dualism as they wrote 'Introduction to statistics' (second edition) in 1972.

Dale J. Poirier decided to develop a course to teach both the classical and

the Bayesian approaches to econometrics in a comparative. M. G. Bulmer introduces statistics using a systematic style, he begins with 'the two concepts of probability' (chapter 1), 'the two laws of probability' (chapter 2) and the pages convey a certain cultural flavor.

Dualists as philosophers

Four people show an evident inclination towards philosophical ruminations. Roy Weatherford writes a book on the philosophical foundations of probability and dedicates 4 paragraphs and the Conclusion of the book to compare various positions. The four paragraphs comment on the positive and the negative sides of the frequentist and the subjective schools respectively.

Philosophical arguments may be found in the book by David Salsburg who spends three paragraphs and other scattered pages on discussing the critical sides of the frequentist and subjective statistics. This book is embellished with several profound considerations.

Ray Hilborn and Mark Mangel follow a special vein of study: they discuss upon the scientific modeling and in Chapter 2 display the manifesto for pluralist models in statistics. Chapter 9 focuses on Bayesian methods and the Epilogue sums up the dualist position of authors.

Fast commentators

Writers discuss the state of the two statistical schools in brief due to the overall style of the book or to other reasons. Larry Wasserman in "A Concise Course in Statistical Inference" sums up strengths and weaknesses of Bayesian inference in Par. 11.9, "[since the book] brings together many of the main ideas in modern statistics in one place".

Wasserman subdivides the discussion of statistical inferences into two distinct parts.

Paul H. Garthwaite, I.T. Jolliffe, and Byron Jones treat the frequentist method in chapters 3-4-5 and the Bayesian in chapters 6-7. They declare that they are not prone to philosophy in the Introduction and make accurate technical comments on the use of the statistical methods in various pages.

Also the books by D. Birkes, Y. Dodge and by U.K. Srivastava, G.V. Shenoy, S.C. Sharma do not devote a special paragraph to discuss the dualist way however they make an accurate account of the statistical methods and add space for reflections. Giuseppe Guerrieri reflects on the subjective probability and devotes a few pages to Bayesianism.

David Williams presents the frequentist and the Bayesian approach as well, emphasizing confidence intervals rather than hypothesis test, and adds Gibbs-

sampling techniques for the practical implementation of Bayesian methods. The author tends to skip philosophical comments and writes:

"Let's see how the Bayesian machine works before getting involved with the philosophy".

Robert Lupton prepares a non pretentious book. He deals with some fundamental topics in statistics which can easily be accessed by the reader. He adds a few comments in the introduction, in the epilogue and in a number of pages.

Moderate Bayesians

Three authors share the Bayesian extraction and even recommend the frequentist tools. Christine R. Jenkins and Jasper Wall describe some topics from classical statistics in Chapter 5 and spend nearly the whole Chapter 3 on an accurate comparison of the schools.

Christian P. Robert co-authored the book "Bayesian Core" (catalogued as a Bayesian work) along with Jean-Michel Marin. We have two editions of the book "The Bayesian Choice": the edition 1994 quotes the frequentism in generic terms and pertains to Group III; and the 2001 edition develops more accurate comments and belongs to Group I. For example Robert presents the classical confidence intervals in Subsection 5.5.2, remarks the least favorable Bayesian answers in Subsection 5.3.5 and recommends the use of Bayesian estimators in Par. 2.4.

Original creators

Andrew B. Gelman and D.A. Nolan write a very original textbook: "Teaching Statistics: A Bag of Tricks" (class ε). The authors see statistics from the perspective of paradoxes and problems which may arise and thus introduce statistics using various agile tools. The authors treat the matter 'by cases' and place the Bayesian and frequentist equations on the same level. The style of the book may be defined as 'asystematic' and thus does not contain a paragraph to methodically compare the Bayesians and frequentist.

David J. Sheskin considers both the relative-frequency and the approaches which he calls 'personalistic'. This original book (class ε) presents 130 statistical tests that pertain to the two schools such as single-sample t-test and Bayesian hypothesis testing. He spends many pages in critical comments.

Harvey Motulsky develops a book on statistics for biologists and doctors. He writes an ample introduction on the Bayesian approach where he dwells on the critical value of statistical prediction. The book is enriched with several comments related to the professional practice.

Group II

Authors belonging to Group I make adequate comments on the complex problems lying at the base of the multiple views of statistics, instead the authors of this group are more concise, and we questioned whether the fundamental definitions lying at the base of the frequentist and the subjective statistics have been given, and how the authors point out the dualist view to the reader.

Balanced Account

We focused on the following basic topics:

1. The Bayes Theorem.
2. The prior/posterior probabilities.
3. The frequentist probability.
4. The subjective probability.
5. The axioms of probability.
6. The large numbers theorem.

It was impossible to make an inventory of other topics because of the variety of concepts treated in Group II. We scrutinized separately the ensuing subsamples from Group II:

$$X = 59 \qquad \text{books from class } \alpha \text{ and } \beta,$$
$$Y = 45 \qquad \text{books from class } \gamma.$$

Table 5 exhibits the approximate percentages of X and Y which provide the formal definitions of the first five topics and the percentages of X and Y which omit those formal definitions.

Table 6 shows how the text defines the large numbers law or assumes the law as given (y), whether the text defines the weak law (w), or the strong law (s), or both the weak and strong laws (ws), or otherwise the authors ignores the law (n).

Results X.1 and X.2 prove how the basic tenets of Bayesianism infiltrate dualist literature. The authors give the frequentist and the subjective definitions of probabilities in equal manner (X.3 and X.4).

More than half of the books do not provide the Kolmogorov axioms in formal terms and the large number theorems (see n in Table 6).

The values of Y are systematically lower than the values of X because the

	The book provides	The book omits
X.1	92	8
Y.1	45	55
X.2	80	20
Y.2	45	55
X.3	71	29
Y.3	27	73
X.4	66	34
Y.4	29	71
X.5	44	56
Y.5	14	86

Table 5

	X	Y
y	22	18
w	10	0
s	0	9
ws	15	4
n	53	69
	100%	100%

Table 6

authors deal with specialist and advanced topics on statistics and assume that the reader is already aware of the basic definitions from 1 to 6.

Remarks

The behavior of authors may be subdivided into four categories which we describe in qualitative terms. We are unable to provide precise statistics because of the contents of this section and the fuzzy boundaries between these categories.

Cat A - Authors present the various mathematical tools and reduce the comments to a minimum. This is the style adopted by pure mathematical

books. For example {27} provides a rigorous account of the frequentist and the Bayesian formulas but does not mention the different meanings of the numbers in terms of objectivism and subjectivism.

Cat B - Authors bring out the weak and the strong sides of each statistical topic treated in the book i.e. estimators, inference methods; but do not make annotations upon the state of the underlying theory/philosophy. e.g.

"The classical approach to point and interval estimation might be considered to have two limitations (...). The Bayesian approach to estimating parameters removes these limitations by formally incorporating our prior knowledge as degrees-of-belief about the value of the parameter and by producing a probability statement about the parameter" {28} p. 27.

Sometimes they add remarks upon the whole statistical school in generic terms. e.g.

"(...) If Bayesian analysis is so great, why have we written a whole book which emphasizes frequentist statistics? We have written about frequentist statistics because Bayesian statistics has a flaw and frequentists (like us) see it as a rather larger one. (...) This is not to say that Bayesian analysis doesn't have a role to play because it is useful in lots of applications. One of them, we have already seen, is diagnostic tests. Another area is in decision making - if we have to decide between two treatments, and the result of our test is not statistically significant, we have to decide on a treatment" {29} p. 319.

Cat C - Authors make comments as those in Cat B do in addition authors declare the different meanings of probability are inessential e.g.

"There is not universal agreement therefore as to the precise definition of probability. We do not have space here to explore the issue further, so we will ignore the problem! The probability of an event occurring will be defined as a certain value and we won't worry about the precise origin or meaning of that value" {30} p. 78.

Cat D - Authors make comments as those in Cat B do, and assert the worthiness of the dualist approach. e.g.

"Antagonism between the two schools is largely a thing of the past and, for the present at least, many statisticians accept that both traditions have something to offer. Whether this continues to be the case once wider experience of the practical, rather than philosophical, differences has been gained remains to be seen. (...) In our view both Bayesian and non-Bayesian methods, have their proper place in statistical methodology" {31} p. 528. e.g.

"Some statisticians are firm adherents of one or other of these schools, but

a widely accepted view is that each has strengths and weaknesses and that one or the other may be preferred in certain context" {32}.

In conclusion the authors explain the use of each Bayesian and each classical formulation in different contexts but tend to skip extended comments on the foundations of statistics and even they declare the irrelevance of basic conceptual differences. This trait basically marks the distance between Group I and II. Authors II highlight the effectiveness of the dualist approach which can surmount a variety of obstacles. Writers show up the capabilities of each mathematical tool in conformity to the purposes of the book, and give the impression that philosophical discussions hamper their efficient illustration of statistical topics. This practice-oriented style contributes to create clear texts but sometimes statistician expressions have the flavor of recipes from a cookbook.

7.1. Group III

There is little to add about the authors of this group. They normally follow the traditional school and add a few notions about the Bayesian methods (see previous description of this group). They absolutely skip the fundamental questions lying at the base of the frequentist and the subjective statistics. Even if they present the subjective probability or the prior/posterior probability, the authors overlook any insight.

8. Discussion and Conclusions

Our study extends and deepens our knowledge about the dualist movement in statistics that turns out to be widely popular within the scientific and the professional community. Dualism spreads over a considerable extent in particular it gained an astonishing success over the past twenty-five years. It seems that the diffusion of Bayesianism favours a more extensive adhesion on the part of dualist statisticians. In fact the dualist production surpasses the pure Bayesian production in the final period of years covered by the sample, whereas significant trends do not emerge during the transition period when the Bayesian school came into the world.

Dualist books treat a broad variety of statistical topics and have different editorial structures as the classes α, β, γ, δ and ε amply show.

Authors reveal different degrees of involvement in dualism. The works in Groups I and II illustrate the achievements of both the schools and constitute

the evident majority (77%), the Group III may be placed at the border line of dualism due to the fast illustration of Bayesianism.

Authors from Group I are concerned with the essential questions of the statistics fundamentals and we have found even philosophical dissertations. At the other hand, the authors from Group II reveal the tendency to overlook the theoretical foundations of statistics; also authors from Group III share this behavior. The greater part of dualist writers (88,6%) are oriented to highlight the professional advantages offered by both the statistical schools and keep apart the foundational arguments since those conflicts can delay and hinder the professional effectiveness. The vast majority of dualists give the impression that the philosophical quarrels interfere with the interpretation and the agile use of statistical methods. Rough speaking dualists suggest putting on different kinds of thinking caps depending on the kind of problems to tackle regardless of the theoretical disputes. Dualist authors adopt formulas deriving from assumptions which clash in point of logic; they do not bother with this incongruence and nonetheless meet mounting success. We believe these threefold circumstances constitute the most significant result of the present research.

References

[1] Anonymous, An international survey of book production during the last decades, Statistical Reports and Studies No. 26/U1289, Bernan Assoc. Publisher (1982).

[2] R.E. Barlow, Introduction to De Finetti, In: *Breakthroughs in Statistics S. Kotz* (Ed. N.Lloyd) Springer (1992), 127-133.

[3] M.J. Bayarri, J.O. Berger, Interplay of Bayesian and frequentist analysis, *Statistical Science*, **19**, No. 1 (2004), 58-80.

[4] G.E.P. Box, An apology for ecumenism in statistics, In: *Scientific Inference, Data Analysis, and Robustness* (Ed-s: G.E.P. Box, T.Leonard, C.F.Wu), Academic Press, Inc. (1983), 51-84.

[5] R. Carnap, *Logical Foundations of Probability*, Univ. Chicago Press (1950).

[6] C. Chatfield, *Problem Solving*, CRC Press (1995), pp. 91-92.

[7] G.A. Crawford, J. Feldt, An analysis of the literature on instruction in academic libraries, *Reference & User Services Quarterly*, **46**, No. 3 (2007), 77-87.

[8] J.D. Emerson, D.A. Colditz, *Use of Statistical Analysis in the New England Journal of Medicine*, MedUseStat (1986), 27-38.

[9] D. Gillies, *Philosophical Theories of Probability*, Routledge (2000).

[10] H. Hakko, P. Rasanen, J. Tiihonen, P. Nieminen, Use of statistical techniques in studies of suicide seasonality, 1970 to 1997, *Suicide Life Threat Behav.*, **32**, No. 2 (2002), 191-208.

[11] J.A. Hokanson, S.G. Bryant, R. Gardner, L.T. Lutmann, G.B. Guernsey, A.C. Bienkowski, Spectrum and frequency of use of statistical techniques in psychiatric journals, *Am. J. Psychiatry*, **143**, No. 9 (1986), 1118-1125.

[12] J.A. Hokanson, D.J. Luttman, G.B. Weiss, Frequency and diversity of use of statistical techniques in oncology journals, *Cancer Treat Rep.*, **70**, No. 5 (1986), 589-593.

[13] J.A. Hokanson, C.M. Stiernberg, M.S. McCraken, F.B. Quinn, The reporting of statistical techniques in otolaryngology journals, *Arch Otolaryngol Head Neck Surg.*, **113**, No. 1 (1987), 45-50.

[14] H. Jeffreys, Russell on probability, *Mind*, **59** (1950), 313-319.

[15] K. Popper, *The Logic of Scientific Discovery*, Hutchinson (1959).

[16] E.W. Shoben, The use of statistics to prove intentional employment discrimination, *Law & Contemp. Probs.*, **46** (1983), 219.

[17] A. Spanos, *Probability Theory and Statistical Inference: Econometric Modeling with Observational Data*, Cambridge University Press (1999).

[18] D.P. Wallace, The use of statistical methods in library and information science, *J. Am. Soc. for Information Science*, **36**, No. 6 (1985).
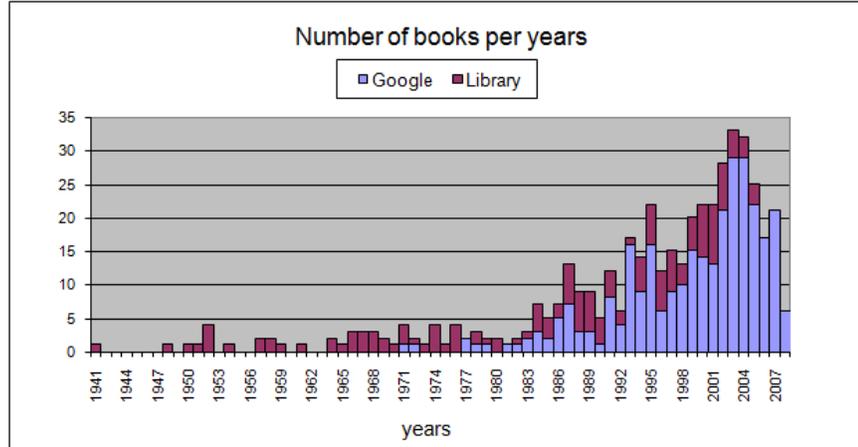
Appendix 1: Figure and Tables

Figure 1: Annual distribution of books from Google and the university library
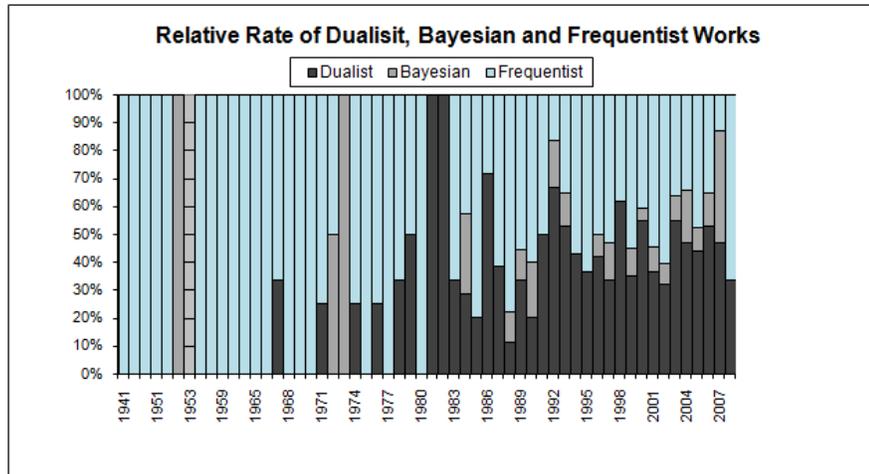

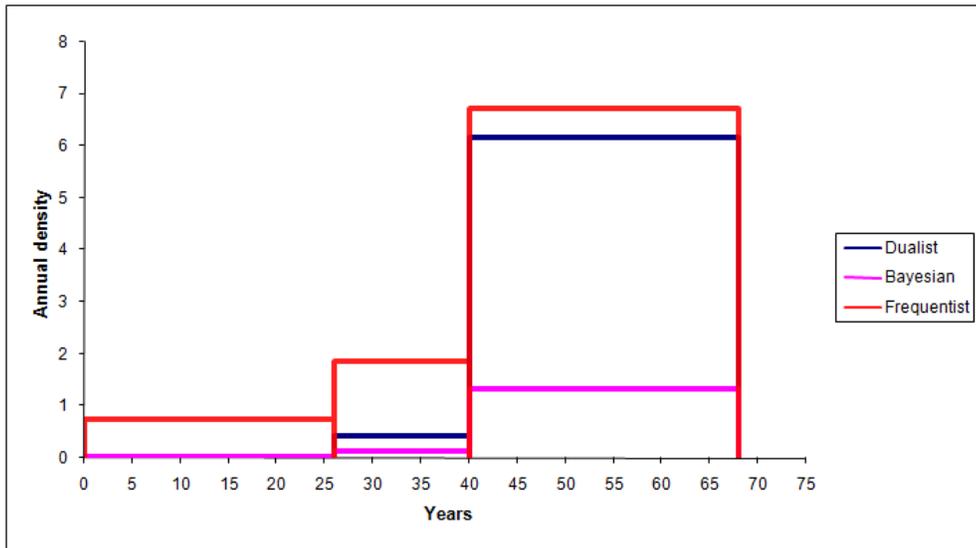Figure 2: Annual distribution of works for each school

Figure 3: Distribution of books over the three periods

Appendix 2: Cited Titles from the Sample

{1} D. A. Kay - *CliffsAP Statistics* - John Wiley and Sons (2004).

{2} A.B. Gelman, D.A. Nolan - *Teaching Statistics: A Bag of Tricks* - Oxford Univ. Press (2002).

{3} D.J. Sheskin - *The Handbook of Parametric and Nonparametric Statistical Procedures* - CRC Press (2003).

{4} M.G. Bulmer - *Principles of Statistics* - Courier Dover Publications (1979).

{5} R. Weatherford - *Philosophical Foundations of Probability Theory?* - Routledge (1982).

{6} U.K.Srivastava, G.V.Shenoy, S.C. Sharma - *Quantitative Techniques for Managerial Decisions: Concepts, Illustrations* - New Age International (1989).

{7} B.H. Wonnacott, R.J. Wonnacott - *Introductory Statistics for Business and*

Economics - John Wiley & Sons (1990).

{8} D. Salsburg - *The Use of Restricted Significance Tests in Clinical Trials?* - Springer (1992).

{9} R. Lupton - *Statistics in Theory and Practice* - Princeton Univ. Press (1993).

{10} D. Birkes, Y. Dodge - *Alternative Methods of Regression* - Interscience (1993).

{11} G. Guerrieri - *Elementi di Statistica* - Cacucci Ed. (1993).

{12} H. Motulsky - *Intuitive Biostatistics* - Oxford Univ. Press (1995).

{13} D.J. Poirier - *Intermediate Statistics and Econometrics: A Comparative Approach* - MIT Press (1995).

{14} C. Chatfield - *Problem Solving: A Statistician's Guide* - CRC Press (1995).

{15} R. Hilborn, M. Mangel - *The Ecological Detective: Confronting Models with Data?* - Princeton Univ. Press (1997).

{16} V. Barnett - *Comparative Statistical Inference* - John Wiley & Sons (1999).

{17} C.P. Robert - *The Bayesian Choice: A Decision-theoretic Motivation* - Springer (2001).

{18} D. Williams - *Weighing the Odds: a Course in Probability and Statistics* - Cambridge Univ. Press (2001).

{19} A.B. Gelman, D.A. Nolan - *Teaching Statistics: A Bag of Tricks* - Oxford Univ. Press (2002).

{20} P. H. Garthwaite, I.T. Jolliffe, B. Jones - *Statistical Inference* - Oxford Univ. Press (2002).

{21} C.R. Jenkins, J. Wall - *Practical Statistics for Astronomers* - Cambridge Univ. Press (2003).

{22} D.J. Sheskin - *The Handbook of Parametric and Nonparametric Statistical Procedures* - CRC Press (2004).

{23} L. Wasserman - *All of Statistics: A Concise Course in Statistical Inference* - Springer (2004).

{24} R.L. Smith, G.A. Young - *Essentials of Statistical Inference* - Cambridge Univ. Press (2005).

{25} R. Russo - *Statistics for the Behavioural Sciences: An Introduction* - Psychology Press (2003).

{26} A. W. Bowman, D. R. Robinson - *Introduction to Statistics: A Computer Illustrated Text* - CRC Press (1987).

{27} G.G. Roussas - *A Course in Mathematical Statistics* - Elsevier (1997).

{28} M.J. Keough - *Experimental Design and Data Analysis for Biologists* - Cambridge University Press (2002).

{29} J. Miles, P. Banyard -*Understanding and Using Statistics in Psychology: A Practical Introduction* - Sage (2007).

{30} M. Barrow - *Statistics for Economics, Accounting and Business Studies* - Pearson Education (2006).

{31} G. Berry, P. Armitage, J. N. S. Matthews - *Statistical Methods in Medical Research* - Blackwell Publishing (2001).

{32} P. Sprent - *Applied Non-Parametric Statistical Methods* - Chapman & Hall (2001).

Frequentist and bayesian statistics: A critique. In Proceedings of the Statistical Problems in Particle Physics, Astrophysics and Cosmology (pp. 3–6). London: Imperial College Press.Google Scholar. Classical and bayesian statistics: A survey upon the pluralist production. International Journal of Pure and Applied Mathematics, 58(3), 255–280.zbMATHGoogle Scholar. Salmon, W. C. (1988).