BIOSTAT at 25

Invited Essays in Theoretical, Biomedical and Social Statistics

Mario Di Bacco, Francesco Scalfari
(*editors*)

Edizioni ETS

www.edizioniets.com

Palazzo Roncioni - Lungarno Mediceo, 16, I-56127 Pisa
info@edizioniets.com
www.edizioniets.com

Distribuzione
Messaggerie Libri SPA
Sede legale: via G. Verdi 8 - 20090 Assago (MI)

Promozione
PDE PROMOZIONE SRL
via Zago 2/2 - 40128 Bologna

ISBN 978-884675409-7

Contents

Foreword

This collection of essays marks the twenty-fifth anniversary of the foundation of the Scuola di Alta Formazione Statistica (International Advanced School on Statistics, aka School of Statistical Inference in Biology and Human Sciences).

In 1993 a summer school was founded to offer researchers in the biomedical sciences, physical anthropologists in particular, an opportunity to further their knowledge of the methods of statistical inference. The first three editions were sponsored by the University of Bologna, and hosted in the wonderful "Villa La Torre" in a small village in the province of Lucca. The villa was bequeathed to the University in 1975 by two Irish-American sisters, Susie e Francis Norton Clarke, to become a "Centro Studi" named after them. The center remained in operation until 1995, then the property fell into disrepair and was eventually sold.

At that point the summer school became itinerant, amid growing economic difficulties (one of the goals set from the start was that fees should be kept as low as possible). We have been guests of several places—Vallombrosa, Cagliari, Rimini, Sulmona and Poppi—supported by a number of local public and private funding bodies; and the continuing success of this initiative is due to both the generosity of our lecturers and the enthusiasm of the participants at large.

In 2000, the school landed in Asti, guest of Asti Studi Superiori (ASTIISS) and supported by the local Fondazione Cassa di Risparmio, and the Piedmontese city has hosted "BiostAT" since then. The continuity offered by the site and the reliable availability of resources have made it possible to expand our activities. On the one hand, the traditional summer courses have been extended to cover the application of statistical inference to all social sciences except economics. On the other hand, other events have been organized, for instance courses of statistical methodology in other universities, and study days that we have called "Colloquia" because based on conversations between statisticians and researchers in the empirical sciences.

In sum, in 25 years of activity we have enjoyed the participation of over 700 students, the majority of which graduate students and young PhDs in biomedical sciences (most of them are today esteemed researchers in their fields) and a number of international lecturers. We wish we could list all their names, and we include them all in a warm "thank you" for contributing with their interest, experience, and friendly collegiality.

About the content of this volume, it suffices to say that the contributions that follow are quite heterogeneous. The range of empirical problems they address and the varied solutions adopted to solve them, in both frequentist and Bayesian inference, simply reflect the attitude that the School had since the beginning: to be exhaustive by touching upon as many examples as possible.

As for the projects that are being planned for the future, we'd rather not say, as projects "are stuff / as dreams are made on".

The Editors.

Prelude

Some reflections on the collaboration between Statisticians and Non-Statisticians

Walter Racugno

University of Cagliari and Polo Universitario of Asti, Italy

racugno@unica.it

Introduction

Throughout the various conferences, workshops and meetings between statisticians and "non-statisticians", meaning researchers from other disciplines who use statistical methods in their studies, there is constant discussion - more and more intense - on the most satisfactory ways of collaboration. In other words, on how to get the maximum benefit for the disciplines involved working together.

The overlay with other sciences is rooted in Statistics since it deals with the study of real phenomena, of any nature, with the aim of dominating the uncertainty with which we perceive them and the uncertainty in measurements or detections that we must carry out to get information on them.

It is certainly no coincidence that several founders of modern Statistics (R. Fisher, C. Gini, B. de Finetti, to name a few of them) were also important researchers in biology, economics, sociology, finance, etc. Their methodological and theoretical contributions to Statistics have often been inspired really by problems posed by those disciplines.

Recently, the request for statistical analyses has grown quickly both because now almost all the international scientific journals require that the experimental data be treated with statistical-mathematical rigor and because in some studies it is possible to have large amounts of data that can be interpreted only through sophisticated statistical techniques.

If it is rare that a statistician does not receive problems from colleagues of other disciplines, on the contrary the presence of specific statistical teachings is increasingly widespread in degree courses, even if not strictly mathematical, or in post-graduate specialization. In view of this, new neologisms have been adopted to indicate this kind of applied studies, e.g. Medical Statistics, Biostatistics, Forensic Statistics, Social Statistics, Economic Statistics, or more generally, Applied Statistics, etc.

Statistics is an applied science, in the sense that even the necessary theoretical developments are aimed at solving real problems and these developments are often motivated by the questions posed by other disciplines.

However, while research in Statistics is having an increasingly rapid evolution towards complex methods, thanks also - as mentioned above - to requests from researchers in other disciplines and to the large amount of data now available, in the "non-statistical field" there is considerable resistance to accepting techniques that deviate from those usually used and already widely known among non-specialists. This divergence between the quality of the problems posed and the resistance to accepting solutions that are not immediately understood represents one of the main difficulties in the scientific collaboration between statisticians and non-statisticians.

Aiming to illustrate this statement I will bring some examples of my personal experiences which are however similar - *mutatis mutandis* - to those of many statisticians. Negletting the surprising and very unusual case of an esteemed dentist who looked for me and a colleague of mine to propose a problem that, in his opinion, was of extreme interest and would have had useful implications in the lives of many people. But we discovered that the talented dentist was also an experienced bridge player who claimed to have built a winning game strategy, that is, in other words, a sequence of *tricks* that led - according to him - to the resolutive game (*manche*). He asked the statisticians to validate his idea and to calculate the probability of success that he expected to be close to 1!

Example 1 - Anaplastic large cell lymphoma

During a long and fruitful collaboration with some oncologists, among the different problems, they asked us to analyze data on children in pediatric age for a study about Anaplastic Large Cell Lymphoma (ALCL), which is a rare cancer disease which affect both children and adults. The aim of the study was to assess the role of the Hsp70 protein (Heat Shock Protein 70 kilodaltons) in association with ALCL. It is therefore a problem of validation of a non-invasive diagnostic test, which is a simple blood test with immediate response, compared to a biopsy.

The dataset consisted on a small sample: 10 patients with ALCL lymphoma (cases) and 4 healthy subjects (controls).

To the statistician, accustomed to believing that the controls are readily available also for rare diseases, it was natural to complain that in this case they were too few. But this request aroused the amazement and also a certain "uneasiness" of oncologists who explained to us that since the certain classification of patients in healthy and sick is done by biopsy, the data on the controls come from "incorrect diagnosis" and certainly not from biopsies performed on healthy subjects. I still remember my mortification when the oncologist, after explaining this to me, asked me: "Do you want more controls?".

What to do in this case? It is evident that with very small samples it is not possible to resort to classical parametric tests, also due to the presence of anomalous data compared to the assumption of normality. Furthermore, the main interest was not to verify if the Hsp70 levels between the two groups were significantly equal or different, but we want to ascertain the validity of the non-invasive diagnostic test compared to a golden test (in this case the biopsy).

We therefore considered the ROC curve, well known in the medical field, for which the area under the curve ($AUC = R$) can be interpreted as a measure of the accuracy of a diagnostic test (see, e.g., Ventura and Racugno, 2017). According to the specialized literature on Hsp70 levels, two independent exponential random variables X and Y were assumed for the protein level in cancer patients and in non-diseased subjects, respectively. Results from a Kolmogorov-Smirnov test supported the choice of the exponential model for these data. Denoting with α and β the rates of the distributions of X and Y, respectively, in this framework the AUC can be written as

$$R = P(X < Y) = \frac{\alpha}{\alpha + \beta}.$$

In view of the small sample size the estimate of R with frequentist techniques, as would be appreciated by medical journal referees, requires sophisticated procedures not shortly explainable (Cortese and Ventura, 2013). However, in this case also the Bayesian approach may present some difficulties. Indeed, the most common choices for the Bayes estimator of R are the mode or the expectation over the posterior distribution $\pi_R(R|x, y)$ which can also be used to derive Bayesian credible sets for R.

First of all, it is required the elicitation of a prior distribution on the complete parameter (α, β), which is difficult both in the subjective and objective Bayesian context. Secondly, cumbersome numerical integration may be necessary in order to derive the marginal distribution of R. This latter difficulty can be avoided using higher-order asymptotics, i.e., accurate approximations of a marginal posterior, which provide very precise inferences on a scalar parameter of interest even when the sample size is small. In this study we used (Ventura and Racugno, 2011) some recent advances in Bayesian inference based on pseudo-likelihood functions, and related matching priors, to perform accurate inference on the parameter of interest R only (Ventura, Cabras and Racugno, 2009). In particular the matching prior have the property to give results where Bayesian and frequentist inference, in the form of posterior quantiles, or confidence intervals, agree to a certain order of approximation.

But on this proposal to solve the problem, there has been a disconnect between statisticians and non-statisticians. Indeed, for the statisticians it deals with procedures based on asymptotic approximations in the Bayesian approach that are quite innovative and of considerable interest for the resolution of problems with small sample size, while the expert of other disciplines has difficulty entering both the methodological and in the interpretation of the results that certainly is far from the reading of the usual *p-value*.

In this specific case, despite overcoming the resistance of the oncologists who had posed the problem, the block was given by the referees of several biomedical journals to which the paper was sent. The referees were able to easily accept the Kolmogorov-Smirnov test and the use of AUC, but already with greater difficulty the Bayesian approach up to severely reject the choice of default priors and the use of pseudo-likelihoods.

The point-by-point answers to the referees were worthless! However the paper, despite some changes tending to illustrate the techniques used as simply as possible, has not found editorial acceptance in any of those journals. Finally, at the end of those exhausting battles it was published (as cited above) in an important statistical journal, reducing the medical problem to a mere application example.

Example 2 - Sexual dimorphism

Another example of successful collaboration muffled by misunderstanding with non-statistical referees, at least as regards the most interesting aspects from the point of view of statistical research, has been with physical anthropologists.

Researches on sexual dimorphism in living populations are generally based on anthropometric measures, that can include the fundamental ones only (stature, weight), or can be detailed into circumferences, diameters, lenghts, and skinfolds.

The more commonly used indices of dimorphism are based on the difference between males and females mean values. However, a dispersion dimorphism can be also present, especially in some measures, such as in skinfolds (Marini *et al.* 1999, 2005). Further, some anthropometric dimensions also show dimorphism of asymmetry: in most cases, length measurements are symmetrically distributed in the two sexes; breadth measurements show a tendency to negative asymmetry in men, and symmetry in women (e.g. biacromial breadth); body circumferences are generally positively skewed, or symmetrical; subcutaneous skinfolds show the greatest sex difference in skewness (Marini *et al.* 2007).

In order to compare, through a synthetic index, the different asymmetry of a certain anthropometric measure between the two sexes, a Skew-Normal model (SN) was adopted and, depending on the case, its variant Skew-*t* (ST) both characterized by a shape parameter λ ($-\infty < \lambda < +\infty$). For $\lambda \to 0$ the SN tends to symmetry and therefore it leads back to the Normal, while the ST to the increase of the degrees of freedom that characterize it tends to the SN.

By limiting these considerations to SN alone, it is important to note that the sign of the parameter λ indicates whether we have a negative or positive asymmetry and its numerical value indicates how far the distribution is from the symmetry. In the study of dimorphism it is therefore essential to estimate the values of λ in the measures of the two sexes to assess whether the asymmetry is significant in each sex and whether it significantly differentiates between males and females.

The classical estimation of λ by the maximum likelihood estimator may present heavy undesirable aspects. Indeed, for small or even moderate sample size, the likelihood can be monotone. Moreover, for $\lambda > 20$, SN models are not very different from each other, because the SN distribution tends to the unique Half Normal distribution as $\lambda \to \infty$, and point estimates cannot be precise, the comparison between different distributions therefore loses practical utility.

To avoid these drawbacks, several modern tecniques have been proposed. In the frequentist approach, some methods are based on an alternative use of a classical bias correction to maximum-likelihood estimation, which anyway never produces satisfying results in general.

In the Bayesian approach, Liseo and Loperfido (2006) show that the Jeffrey's prior of λ is proper, a particular situation for this class of priors, given that the range of λ is unbounded. Hence the posterior distribution for λ is always a proper distribution and its mode or median produces finite estimates, which are shown to have good frequentist properties.

Later, Cabras *et al.* (2012) have proposed an approach of performing a default Bayesian analysis on the shape parameter λ of the SN based on a suitable pseudo-likelihood function and a matching prior distribution for this parameter, when location and scale parameters are unknown. This approach is important for both theoretical and practical reasons. From a theoretical perspective, the proposed matching prior is proper thus inducing a proper posterior distribution for the shape parameter, also when the likelihood is monotone. From the practical perspective, this approach has the advantages of avoiding the elicitation on the location and scale parameters and the computation of multidimensional integrals.

The aim of the examples presented want to illustrate two different problems that typically arise in the collaboration between statisticians and non-statisticians. In the first one the statistician has found in the applicative problem posed by the oncologist the idea to develop some theoretical and applicative aspects of the statistical methodology; in the second one the statistician used advanced theoretical developments for a challenging application in the anthropological field.

In both cases the proposals were not appreciated by the referees who would have preferred to be in front of consolidated statistical techniques, but ineffective to solve those problems.

Example 3 - Independence of two qualitative variables

A last case different from the previous ones, but perhaps even more disconcerting, can be described through a fictitious example, but conceptually similar to that posed by a hematologist.

It is well known and generally accepted that the blood group in mankind is not associated with sex, in other words, belonging to group A, B, AB, 0 does not depend on being male or female.

The hematologist has collected more than 3000 blood samples detecting sex and blood group of each subject and, aiming to confirm the non-association between the two qualitative variables, he calculates the Chi-square of the 2×4 table. Due to the high sample size he finds a very large Chi-square value that leads to the rejection of the hypothesis of independence between blood group and sex.

The hematologist asks for assistance to the statistician.

I invite readers to pose this question to different statisticians: in light of over 3000 data observed, how can we show that there is actually no association between the two variables? Certainly different solutions will be proposed, some even very complex, someone will propose to use descriptive methods for large amounts of data (*Big Data*), someone else will propose to dissect the analysis in different steps (for example studying group by blood group and then ...), still others to use generalized linear models ...

In short, it will be very difficult to counter the trust that the hematologist has always placed in the Chi-square!

Short conclusion

I believe it is important to intensify the relationships between statisticians and non-statisticians, extending them beyond the usual requests for data analysis at the end of the experiment, but starting the collaboration from the research design and in particular to planning the data collection. It is also essential to develop mutual updates that allow greater agility in the linguistic exchange. Indeed, in the early stages of each interdisciplinary collaboration the first difficulty to overcome concerns precisely the understanding of the specific problem that the expert poses to the statistician and its consequent formalization, while the non-statistician has difficulty interpreting the technical or methodological solutions that the statistician proposes.

From this point of view, events like those that BioStat (School of Inference Statistics in Biology and Humanities, founded in 1993 and still directed by M. Di Bacco at the University of Asti) proposes for many years are moving in this direction.

References

Azzalini, A. (1985). A class of distribution which includes the normal ones. *Scand. J. Stat.*, **12**, 171–178.

Cabras, S, Castellanos, M.E., Racugno, W., Ventura, L. (2012). A matching prior for the shape parameter of the skew-normal distribution. *Scand. J. Stat.*, **39**, 236–247.

Cortese G., Ventura L. (2013). Accurate higher-order likelihood inference on $P(Y < X)$.Ê*Computational Statistics*, **28**,1035–1059.

Liseo, B., Loperfido, N. (2006). A note on reference priors for the scalar skewnormal distribution. *J. Stat. Plan. Infer.*, **136**, 373–389.

Marini, E., Rebato, E., Racugno, W., Buffa, R., Salces, I., Borgognini-Tarli, M.S. (2005). Dispersion dimorphism in human populations. *American J. Phsycal Anthropology.*, **127**, 342–350.

Marini, E., Cabras, S., Rebato, E., Buffa, R., Salces, I., Borgognini-Tarli, M.S. (2007). Sex differences in skinfold variability across human populations and during the life cycle. *Ann. Hum. Biol.*, **34**, 377–392.

Ventura, L, Cabras, S, Racugno, W. (2009). Prior distributions from pseudo-likelihoods in the presence of nuisance parameters. *J. American Stat. Ass.*, **104**, 768–774.

Ventura, L, Racugno, W. (2011). Recent advances on Bayesian inference for $P(X < Y)$. *Bayesian Analysis*, **6**, 1–18.

Ventura, L, Racugno, W. (2017). *Biostatistica. Casi di Studio in R.* Egea Ed.

Skulls and transvariation

Gianmarco Altoè [1] Giuseppe D'Amore [2] Francesco Scalfari [2]

[1]Dipartimento di Psicologia dello Sviluppo e della Socializzazione, Università degli Studi di Padova, Italy. Email: gianmarco.altoe@unipd.it

[2]Biostat Scuola di Alta Formazione Statistica, Asti, Italy

Abstract

The reconstruction of the history of human populations is made possible by the application of different approaches. In Physical Anthropology, the analysis of metric characters of the skull is fundamental and regularly applied to determine the biological relationships between populations alongside archaeological, genetic, and linguistic information. The purpose of this study is to propose a new index (D_A) to evaluate the distance between pairs of skull populations based on a multivariate set of craniometric variables. The new measure reflects the concept of transvariation area developed by the Italian statistician Corrado Gini in 1916. As an illustrative example, an application to a practical case will be presented. Specifically, an expert anthropologist is asked to identify 3 craniometric variables and to assign the respective weights in terms of importance to evaluate the distance between pairs of 9 groups of skulls with different geographical origin and ethnic affiliation. The results indicate that D_A is an efficient measure to evaluate and summarize the distance between groups. Moreover, due to its immediate geometric interpretation, it can easily be interpreted graphically. The usefulness of this new measure in Physical Anthropology, its natural extension to other scientific domains, and other potential applications are discussed.

Keywords and phrases. transvariation area, craniometric variables, distance indices

"Les particularités physiques, intellectuelles, morales, qui distinguent les groupes humains disséminés à la surface du globe, accusent-elles entre ces groupes des différences radicales? ou bien, malgré les apparences contraires, l'homme est-il partout le même au fond? En d'autres termes, existe-t-il une seule espèces d'hommes? Cette question est toute moderne. [...] Du peu que ont écrit sur cette matière, on doit conclure quà leurs yeux la nature de l'homme est partout la même, et que des conditions extérieures, le froid et la chaleur particulièrement, font seules varier ses caractères physiques."

"Do the physical, intellectual, and moral peculiarities which distinguish human groups scattered on the surface of the globe, show radical differences between these groups? or, despite appearances to the contrary, is man, after all, the same everywhere? In other words, is there only one species of men? This question is totally modern. [...] From the little that has been written on this subject, we must conclude that in their eyes the nature of man is everywhere the same, and that external conditions, in particular the cold and the heat, only changes his physical characters."

Increase limit in top speed of hundred meters' track competition

L. Bollani

ESOMAS Department, University of Turin, Italy

luigi.bollani@unito.it

Abstract

In this article, a logistic model is introduced to adapt to the speed data obtained in the modern Olympics for the 100-meter men's racing competitions. Firstly, a gamma truncated distribution is proposed to a priori describe the maximum reachable speed, on the basis of observations collected in literature. Secondly, a posterior distribution is derived considering the likelihood expressed by the data, with respect to the logistic model originally introduced. Some considerations are made comparing prior and posterior distributions.

Keywords and phrases. logistic model, gamma distribution, Bayesian probability model, prior and posterior distributions.

1 Introduction

The interest in the field of predicting sporting results at the limits of human possibilities often occurs in many disciplines and is made spontaneous by the atavistic interest of man in comparison with his own limits. On the other hand, the history of sport teaches us to doubt the physiological hypotheses aimed at proposing impassable thresholds. See, for example, the 1945 Roger Bannister feat that knocked down the four-minute wall to cover a mile or the result of Enzo Maiorca in 1962, which exceeded fifty meters of depth. One way to tackle the problem from a rational point of view is to quantitatively express an opinion on a variable threshold, attributing to it a probability distribution congruent with current knowledge and with the sports records so far achieved in a given discipline, reserving the right to update this opinion if the conditions will vary over time. This is what will be proposed in this paper, referring to the hundred-meter speed records on the men's track in the modern Olympics, with the use of a Bayesian probability model (Daboni L., Wedlin A.,1982; de Groot M.H., Schervish

Weighted Likelihood methods for Data Reduction

Luca Greco

Department DEMM, University of Sannio, Italy

luca.greco@unisannio.it

Abstract

Standard data reduction techniques, such as principal component analysis, discriminant analysis, cluster analysis, exhibit lack of robustness with respect to the occurrence of outliers, anomalous values that can completely break down classical procedures, hence leading to unreliable conclusions. This unpleasant behavior stems from the fact that they rely on the sample mean vector and sample covariance matrix. Then, robust data reduction methods can be defined by supplying robust estimates of multivariate location and scatter. Furthermore, formal rules for the purpose of outlier detection can be obtained. In particular, the interest focuses on those techniques driven by the employ of weighted likelihood multivariate estimates. The behavior of several weighted likelihood based techniques has been investigated by some real data examples.

Keywords and phrases. Cluster analysis Discriminant Analysis Mahalanobis distance Multivariate Normal Outlier detection Pearson residuals Principal component analysis Robustness Weighted Likelihood

*Mathematics Subject Classification (2000).*MSC 62F35 MSC 62G35 MSC 62H25 MSC 62H30

1 Introduction

Data reduction is the process of summarizing the data by aggregating information. This task can be achieved by reducing the number of dimensions, the number of observations or even both at the same time. Principal component analysis (PCA) is a very popular method aiming at extracting features from the data at hand, whose dimensionality is remarkably lower than that of the original data, while retaining the most part of its variability. Cluster analysis (CA) is an unsupervised sample reduction technique that allows to aggregate sample units leading to very few cluster profiles. Discriminant analysis (DA) carries sample reduction as well, but in a supervised fashion.

Many multivariate techniques often rely in the assumption of multivariate normality. Actually, the sample mean vector and covariance matrix provide a very simple description of the overall shape of the data by the related ellipsoids (Huber and Ronchetti,

Multi-one-sided tests for multivariate ordered categorical data with medical applications

Fortunato Pesarin

Department of Statistical Sciences, University of Padova, Italy

fortunato.pesarin@unipd.it

Abstract

This article deals with permutation testing methods for univariate and multivariate ordered categorical data. The specific interest is on testing for stochastic dominance, i.e. for restricted alternatives. Several solutions to the univariate case based on restricted maximum likelihood ratio tests have been proposed in the literature. These solutions are generally criticized since their asymptotic null distributions are mixtures of chi-squared variables with weights depending on the unknown population distribution F and so the related accuracy is difficult to assess. Further, testing for stochastic dominance in multivariate cases by likelihood approach is known as an even more difficult problem. By working within the conditioning on a set of sufficient statistics in the null hypothesis and the nonparametric combination of dependent permutation tests it is possible to find exact solutions to problems of that kind. Solutions for two-sample designs guided by two medical application examples are provided.

Keywords and phrases. Conditional inference, Multivariate permutation testing, Nonparametric combination, Restricted alternatives, Stochastic dominance

1 Introduction

Problems of testing with ordered categorical variables are frequently met in many disciplines: biostatistics, clinical trials, genetics, marketing, pharmacostatistics, psychology, quality control, social sciences, technology, and so on, where a finite number of $V \geq 1$ of such variables are observed on each individual unit. Testing of hypotheses with ordered categorical variables is known to be quite a difficult problem when testing for stochastic dominance, that is for a set of restricted alternatives. Stochastic dominance problems are of specific interest in application contexts since are frequently encountered in practice and present peculiar difficulties especially within the framework of likelihood ratio

Longitudinal analysis in the medical sciences: an application to the investigation of post nephrectomy course

Sergio Pezzulli and Rodolfo Rivera

Kingston University London
and
Nephrology and Dialysis Division, San Gerardo and Desio Hospital, ASST, Monza

Abstract. We present the initial results of a retrospective study conducted on 245 patients who underwent unilateral Radical Nephrectomy (RN) as treatment of an urological pathology. The recorded postoperative creatinine levels are analysed for determining the principal patterns of variation in the two year following the intervention. The observational data, which are sparse and right censored, are linearly interpolated and then analysed via a robust form of Principal Component Analysis (PCA). The results seem to indicate that the quantitative description of the main patterns can be reduced to few descriptive measures based on postoperative observations.

Keywords: Observational studies, Radical Nephrectomy, Principal Component Analysis, Functional Data Analysis, Robustness

1 Introduction

The continuous developments of experimental techniques in Biology, Physiology and Medical Sciences often provide more abundant and structured data that may give newer insights into biological processes. In order to understand these processes and their dynamic interaction, we need statistical tools that can help to identify, interpret and use these information.

Often experimental data describe the dynamic behaviour of an individual process observed in time. The systematic approach for answering the questions posed by time dependent observations is commonly referred to as time series analysis [1],[2]. Time series are collected in the social and the health fields in order to study the dynamic of birth rates [3], mortality, school enrolment, vaccine immunization [4], etc. In the epidemiological area, applications may aim at determining number and characteristics of influenza cases observed over some time period, or the effects of the most common children's virus exanthemas infections after vaccination [5].

Medical data are often congested with noise, which may represents biological variability, effects of interventions performed or a variety of external factors that cannot be controlled for. Statistical methods can help us to reliably interpret

Statistical Analysis for Case-Control Studies

Tiberiu Postelnicu

Commission for Biometrics, Romanian Academy

tposteln15 @yahoo.com

Abstract

The aim of this paper is to present some new aspects regarding statistical methods for case-control studies. Identification of factor that increase the chances of a certain disease is one of the classical and central issues in epidemiology. In this context, the odds-ratio, as typical measure of association between disease and risk factor is presented for matched studies, as well as using Bayesian inference. The use of complex stratified sampling procedures and of the meta-analysis are discussed.

1. Introduction

Events with a variability in attributes or responses are not predictable, and there need statistical methods. If we consider two or more different events which are *independent* of each other, then to get the probability of a combination of specific outcomes for each of the events we must multiply the individual probabilities of those outcomes. By independent we mean that if we know the outcome of one event, this tells us nothing about the other event. If two events are not independent, the multiplicative property does not apply. The idea is used in reverse in case, of uncertainty to investigate whether two events are independent. For example, in a *case-control study* patients with a disease (cases) are compared with people without the disease (controls) with respect to some possibly hazardous exposure earlier in their life [1].

Measuring consciousness and assessing language in epilepsy: application of statistics to neuroscience

Cristina Rosazza

Neuroradiology Dept., Fondazione IRCCS Istituto Neurologico Carlo Besta Milano, Italy

cristina.rosazza@istituto-besta.it

Abstract

The use of statistical analysis is crucial in neuroscience and in particular in neuroimaging studies, where a huge amount of data is obtained and the need to make accurate diagnosis and prognosis is strong. Statistics plays a critical role in understanding the data and obtaining relevant results that can be used and interpreted by neuroscientists. In this chapter I will illustrate how statistical analysis has been successfully applied to two neuroimaging researches, conducted mainly with Magnetic Resonance Imaging (MRI). The first one concerns Disorders of Consciousness, a troublesome spectrum of conditions comprising vegetative state and minimally conscious state. Statistics has been useful to understand whether different neuroimaging techniques can distinguish the neurological conditions and to obtain the corresponding diagnostic accuracy. The second research concerned Epilepsy, a neurological condition that can be treated surgically when drugs do not control seizures. Statistics has been critical to understand whether patients have a defective language performance compared to healthy subjects before surgery and whether functional MRI can be useful to identify the patients at risk of postoperative language deficits.

Keywords and phrases: Disorders of Consciousness (DOC), Magnetic Resonance Imaging (MRI), Temporal Lobe Epilepsy (TLE), logistic regression, receiver operating characteristic (ROC) curves, Q test, independent-sample t-tests, Mann-Whitney U tests, Pearson correlations, linear regression analyses.

1 Introduction

Neuroscience is an interdisciplinary field that aims to study the brain and behavior. It includes many disciplines like biology, medicine, psychology but also statistics, physics and engineering. In the clinical setting, statistics plays a key role because it allows the definition of a diagnosis, a prognosis, and is useful to evaluate the effectiveness of treatments. Statistics is particularly useful in neuroimaging where a typical Magnetic

Can a Witch teach Stats?

Monica Bailot, Rina Camporese, Silvia Da Valle, Sara Letardi, Monica Novielli, Susi Osti

Department for data collection and development of methods and technologies for the production and dissemination of statistical information,
Italian national institute of Statistics, Italy

bailot@istat.it campores@istat.it davalle@istat.it letardi@istat.it novielli@istat.it osti@istat.it

La conoscenza non è
una quantità,
è una ricerca.(G. Rodari)

Abstract

Regardless of calculus ability, children need to approach statistics and stochastic literacy as soon as possible, so as to build up their ability to deal with uncertainty when making judgements and decisions. Moreover, statistics and probability are mandatory in curricula starting from primary school. Authors present a didactic project in which they explore the possibility of spreading statistical culture through the methods of fantastic narration, choosing an unconventional language to communicate statistics to kids, not forgetting methodological rigor anyway.

Key words: Kids, statistics, probability, fables

1 Introduction

In today's digital age, where everybody is continuously exposed to an overwhelming wave of data and information, it is crucial to prepare future generations to navigate the world of statistics. It becomes increasingly necessary to educate young people to read the reality through a critical interpretation of data and to promote the ability to manage uncertainty.

Otherwise, there is a high risk that younger generations, becoming adult citizens, would not be able to discriminate between credible and false information. In fact, they could be unable to critically interpret the messages they might be confronted with (Sharma, 2017).

The Italian Ministry of Education underlines how "statistics uses mathematics to explain phenomena and trends of our nature, world and society".

Statistics can therefore be used as an effective "Trojan horse" to bring students closer to mathematics and to its powerful ability to explain and interpret the world with a critical spirit, using data to support their opinions (MIUR 2018).

THE MEASURMENT OF WELL-BEING IN THE WORKPLACE: A FACTOR EXPLORATORY ANALYSIS

A Bitetto M.D. Ph.d.
Department of Mental Health, Asst Monza, Italy
Department of Management, University of Torino, Italy

Abstract

Well-being is a positive outcome measure that includes global judgments of life satisfaction and feelings of joy (Lyubomirsky S, et al. 2005). Good living conditions in personal relationships, housing and employment are fundamental to well-being. In work place well-being influences many aspects as worker health and safety but also productivity (Majer V. 2009; Idris M. A. et al., 2012). Therefore its tracking in surveys on job conditions is important for work policy even if is difficult to measure. The main objective indices used are the days of paid sick leave and work accidents. However, well-being is a subjective multidimensional variable, typically measured by self-report (Kahneman D et al., 2004). Currently there is a general agreement to use both objective and subjective measures, with the suggestion to in-depth factor exploratory analysis of data collected by questionnaires. Following, as example of application, are reported the results of a survey on effects of job well-being, conducted in a Milanese hospital between 2007 and 2010.

Keywords and phrases. well-being, sick leave, work accidents, work safety, latent variable, objective and subjective variables, factor explanatory analysis

Introduction

There is no consensus around a single definition of well-being: at minimum, well-being includes the presence of positive emotions and moods (e.g., contentment, happiness), the absence of negative emotions (e.g., depression, anxiety), satisfaction with life and positive functioning (Lyubomirsky S, et al. 2005; Diener E, et al.2004).

While much of the stress that people experienced, comes from their personal lives (Sadock & Sadock, 2003) job conditions can make it worse or sometimes cause discomfort with permanent damage at the worker health (Marri G. & Oddone I. 1967; Kendler 2003; Zwetsloot G. et al, 2017).

Moreover in the workplace, employee well-being influences certain behaviours such as productivity, relationships between colleagues and superiors and, in general, allows a better climate to be achieved (Lazarus R.S. & Folkman S. 1984, Majer V. 2009).

Much research has been conducted on this topic over the past 50 years but findings are far from conclusive as a consequence of the difficulties in adequately measuring such a complex phenomenon.

The most important objective index of well-being in the workplace is the days of paid sick leave. A report of WHO in 2010 (Scheil-Adlung X. & Sandner L. 2010) underlies how paid sick leave plays a crucial role especially in times of crises where many workers fear dismissal and discrimination when reporting sick. In fact, the absence of paid sick days forces ill workers to decide between caring

Bayesian inference in forensic science

Silvia Bozza

Department of Economics, Ca' Foscari University of Venice, Italy

School of Criminal Justice, University of Lausanne

silvia.bozza@unive.it; silvia.bozza@unil.ch

Abstract

Forensic scientists deal with the evaluation of a link between recovered material of unknown source found at a crime scene and control material coming from a suspect. The assessment of the value of the scientific evidence is typically performed by means of a likelihood ratio, a well established metric in forensic science. However, the derivation of a likelihood ratio may represent a demanding task with several sources of uncertainty, and this has originated a large debate about what should be the most appropriate way to take charge of such uncertainty while presenting expressions of evidential value at trial. In such a context, Bayesian networks represent a powerful tool that can be used to study, develop and implement probabilistic procedures for evaluating the probability value of the scientific evidence in forensic science or of an hypothesis of judicial interest.

Keywords. Bayes theorem, Bayesian networks, evaluation of evidence, discrimination, forensic science, likelihood ratio, uncertainty.

1 Introduction

Forensic science relies on a body of scientific principles and practices to help with issues in legal proceedings, such as criminal investigations. In many circumstances, forensic scientists are asked to help reconstructing the dynamic of past events of some judicial relevance (e.g., the dynamic of a murder). The natural response to face the unavoidable uncertainty associated to past events (e.g. a murder) is the collection and sound use of information, typically data resulting from comparative analyses of evidential material (comparative in the sense that some characteristics of traces recovered at the crime scene are compared to those of material originating from a given suspect). The assessment of the evidential strength of analytical results within the event under investigation may certainly represent a fundamental step in the interpretation of facts. Although uncertainty can not be eliminated, it can be measured. The results of comparative analyses

A latent Markov model for evaluation of partnership dynamics in young American men

Alessio Farcomeni

Department of Public Health and Infectious Diseases Sapienza - University of Rome, Italy

alessio.farcomeni@uniroma1.it

Antonella Cuteri and Gabriella Cuteri

Department of Statistics Sapienza - University of Rome, Italy

Abstract

We evaluate the impact of some economic, social and demographic factors on marital status evolution. We focus on a panel of American men aged between 18 and 35 years. We model marital status (married/cohabitant vs otherwise) through a recently introduced latent Markov approach. In doing so, we adjust for unobserved heterogeneity by modeling the random intercept as a first-order homogeneous Markov chain. The state dependence effect is estimated simply by including the lagged response variable among the covariates. We find a strong state dependence for marital status, and a strong effect of employment status both in the current and previous year. It can be concluded that a stable employment and income source may favour stable union formation.

Keywords and phrases. Employment, Marital status, latent Markov model, Panel Study of Income Dynamics, State dependence

1 Introduction

At the beginning of adulthood every individual usually makes several different choices about his/her own life. Many important choices concern leaving the parent's house, dealing with relationships, leading a working life, etc. Couple membership roles may remain stable, or they may change by entry into marriage/cohabitation or dissolution of marriage/cohabitation. This kind of dynamics is well known to affect and be affected by economic well-being, employment status and health conditions. Many studies focus on economic well-being following a couple disruption, with particular attention to gender differences. There are two different points of view: part of existing literature supports that women experiencing couple breakdown undergo a worsening

Bayesian Statistics by Example: A Simple Meta-Analysis of Parapsychology Data

Wesley O. Johnson and Jessica M. Utts
Department of Statistics, University of California, Irvine, CA, USA
Irvine, CA 92697-1250 USA

June 29, 2018

Abstract

Although statisticians have the job of making conclusions based on data, for many questions in science and society prior beliefs are strong and may take precedence over data when people make decisions. For other questions, there are experts who could shed light on the situation that may not be captured with available data. One of the appealing aspects of Bayesian statistics is that the methods allow prior beliefs and expert knowledge to be incorporated into the analysis along with the data. One domain where beliefs are almost sure to have a role is in the evaluation of scientific data for extrasensory perception (ESP). Experiments to test ESP are often Binomial, and they have a clear null hypothesis, so they are an excellent way to illustrate hypothesis testing. Incorporating beliefs makes them an excellent example for the use of Bayesian analysis as well. In this paper, data from one type of ESP study are analyzed.

1 Introduction

This paper presents a simple Bayesian analysis to address a question that has fascinated humankind for centuries, for which there is still no definitive answer, and for which there is now enough data to make an intriguing statistical study. The question is whether it is possible for people to get information using methods commonly called psychic abilities or extrasensory perception (ESP). The second author has done extensive work with parapsychologists, scientists who use laboratory studies to investigate whether psychic functioning is possible, and

Statistical issues in Bayesian cost–effectiveness analysis

Elías Moreno, F.J. Vázquez–Polo,
M.A. Negrín and M. Martel–Escobar

E. Moreno, Dpt. of Statistics, University of Granada, Spain

F.J. Vázquez–Polo, M.A. Negrín and M. Martel–Escobar,

Dpt. of Quantitative Methods, University of Las Palmas de Gran Canaria, Spain

emoreno@ugr.es,

{francisco.vazquezpolo or miguel.negrin or maria.martel}@ulpgc.es

Abstract

Cost–effectiveness analysis of medical treatments search for choosing an "optimal" treatment among a set of $k \geq 2$ alternative treatments $T_1, ..., T_k$ for a given disease. It is imposed that the cost and the effectiveness of the treatments are taken into account in the selection procedure.

We focus the problem as a Bayesian statistical decision problem, present their elements and illustrate the procedure. Further, we discuss some difficulties arising in cost–effectiveness analysis when heterogeneity is present in the cost and effectiveness data. Heterogenous data implies in cost–effectiveness analysis the need of considering special statistical techniques such as Bayesian meta–analysis and Bayesian probabilistic clustering.

Keywords and phrases. Cost and effectiveness of a treatment, predictive reward distribution of a treatment, optimal treatment, utility function.

1 Introduction

Health Economics is an area of the field of Economics with an intensive recent development. The major concerns of researchers in this area is the comparison between medical treatments based on their effectiveness and cost. It is accepted that health resources are limited and effectiveness comes at a price. As control over health expenditure has increased over the last thirty years, the term cost-effectiveness (CEA) has gained in popularity.

Statistical methods used to evaluate perception of risk(s)

Mariana Toma-Drane, PhD, MPH, MIS
CEO, Health Assessment & Risk Management, USA

The purpose of this paper is to look at statistical methods that are successful in identifying and capturing long-term perceived risk after environmental accidents among exposed individuals.

Two series of mixed linear regression models were estimated to assess long-term perceived risk among families within the exposed community. First one assessed if relocation away from the contaminated area(s) improved positively subjects perception of being at risk. The second one introduced the interaction between the concepts of relocation among the individuals within the family.

It was observed that relocation was an underlining factor in the level of perceived risk. The higher the risk was perceived the higher the likelihood of relocation. While taking this decision was not easy especially for women, statistic shows that it was regarded as parents' protective mechanism for their offspring.

Key words: health risk, perceived risk, relocation.

• Introduction

In general, risk perception is a highly personal process of decision making, based on an individual's frame of reference developed over a lifetime. When it comes to making decisions about health and safety, individuals have different perspective and don't always worry the most about the greatest real threats (Slovic P, 1987 & Fischhoff B, et al., 1978). Usually, there are numerous hazards to consider and for each one of them there are many possible scenarios that could unfold depending on timing, magnitude, type and location of the hazard.

It was suggested that risk is mentally assessed in similar ways by individuals, while risk perception is shaped by several mostly unconscious emotional processes. Previous studies found that the human brain is designed to react quickly and defensively to perceived threats of any kind (LeDoux J., 2012) i.e., physical threats, sights, sounds, smells, and just words or memories associated with fear or danger (Starr G, et al. , 2000).

Latest advances on Objective Bayesian model selection for survival regression

Stefano Cabras[(1,3)], María Eugenia[(2,3)] Castellanos and Silvia Perra[(3)]

[(1)] *Universidad Carlos III de Madrid (Spain)*

[(2)] *Universidad Rey Juan Carlos (Spain)*

[(3)] *Università di Cagliari (Italy)*

stefano.cabras@uc3m.es, maria.castellanos@urjc.es, silviaperra@gmail.com

Abstract

In this article view review some recent advances on Objective Bayesian model selection for regression model with survival data. Such a topic is of particular importance in Biostatistics. This article tries to condensate into few pages this large and debated topic in Bayesian statistics, afforded under the Objective Bayes point of view in which priors are derived from formal rules. Here we review how to derive such priors for producing consistent regression model selection procedures when the sample are differentially informative due to the presence of censored observations.

1 Introduction

We consider the problem of model selection in regression analysis when the response variable, Y, follows a parametric distribution (e.g. generalized gamma, Weibull, log-normal, exponential, etc.) and observations are right censored, with known probability distribution on the hidden censoring variable. For model selection we use Bayes Factors (BFs) (Kass and Raftery, 1995) which are unscaled when improper priors are used. To overcome this problem and to approximate the underlying BF, we recur to the notion of Fractional (FBF) and Intrinsic BFs (IBF) (Berger and Pericchi, 1996; O'Hagan, 1995), which have been proposed in the literature. Both depend on the size, N_t, of the minimal training sample (MTS) and in particular the IBF also depends on the specific MTSs used. In the context of censored variables, N_t is a random variable, whose distribution is specified below. For this reason, the IBF is non-parametrically

A Bayesian Approach to Comparing Several Groups or Treatments as Used in ANOVA

John Deely

Abstract

This paper discusses the Bayesian approach as an alternative to some classical ANOVA procedures. Two specific data sets are used to illustrate the Bayesian approach and results. Posterior probabilities for which unknown mean is largest among the group of means and by how much are computed. The value "how much" is used to talk about "practical significance". The concept of Bayesian 'interaction' is also described.

1. Introduction

The purpose of this paper is to illustrate the value of a Bayesian approach in practical problems involving three or more groups or treatments which are to be compared. This will include application to ANOVA models. The reader is referred to Christensen, etc. (2011) for ANOVA ideas.Using practical criteria for comparisons, the advantage of the Bayesian approach will be demonstrated in two situations.

Situation one, will deal with an experiment consisting of lamb data from five diets and the desire is to find the diet with the largest mean weight gain.

Situation two will deal with data consisting of bushels of corn per acre from four varieties of corn and three varieties of fertilizer and the desire is to determine if a particular fertilizer interacts with a particular variety of corn to give the largest mean production.

Both of these problems lend themselves in a very practical way to a Bayesian analysis. That is, there is a likelihood function that describes the data's relationship to some quantities of interest labeled the population parameters. These parameters of interest will then be given a prior distribution, which will depend on the type of prior information available in any given situation. Then using Bayes theorem the posterior distribution of the parameters of interest can be computed and using the posterior, various useful and important probabilities can be computed. These will be discussed after introducing the data sets. It is to be noted that this posterior distribution is always conditional upon the data that has been observed. It does not suppose that imaginable data sets have occurred and that probabilities of imagined data have been computed for various values of these parameters. These statements will be illustrated specifically in the two problems discussed in this paper.

In Section 2 we first of all introduce a data set concerning a study of five diets for lambs. In Section 3 we develop the formulas and notation for the Bayesian analysis. The classical frequentist analysis of that data is discussed and the aspects to be addressed by a Bayesian analysis are mentioned. In Section 4 various Bayesian approaches are introduced and computations displayed. A concluding discussion of these results is given in Section 5.

Bayes' rule and optimal updating

Viviana Doldi
University of Pavia, Italy
Mario Di Bacco
Polo Universitario of Asti, Italy

Abstract

Bayes' rule is used in order to update probabilities, because of the joint probability theorem and the coherence principle proposed by de Finetti. In this paper we prove that if updating is necessary (that is, if someone prefers not to be satisfied by his first evaluation over a trueness of an event) then Bayes' rule is the only way to update probabilities.

AMS (2000) subject classification: 62B10, 62C10
Keywords: Updating, Expectation, Bayes's Rule.

1 Introduction

Let H be an event. We say $H = 1$ if H is true, that is happened, and $H = 0$ if H is false, not happened.
The Decision Maker DM measures his trust in the trueness of H, in $H = 1$, by the probability p and plans an experiment $\begin{pmatrix} X \\ \Pi \end{pmatrix}$, having results

$$X = \{x_1, \ldots, x_i, \ldots, x_n\}, \quad 1 < n < +\infty; \tag{1}$$

he assigns, as a measure of his confidence, probability $\pi_i > 0$ to the event "$\begin{pmatrix} X \\ \Pi \end{pmatrix}$ will finish in x_i", being $\sum \pi_i = 1$.
DM has planned $\begin{pmatrix} X \\ \Pi \end{pmatrix}$ in order to update his confidence in $H = 1$, that is, to update p. For this purpose he will use Bayes' rule: if he observes $X = x_i$, he will update p by

$$B(p, L(x_i), \pi_i) = \frac{p \cdot L(x_i)}{\pi_i}, \tag{2}$$

Whither Bayesian Statistics ?

Frank Lad

Department of Mathematics and Statistics, University of Canterbury

Abstract

As a tribute to the inquisitive and committed students in many cohorts I have met at the Asti Summer School in Statistics, I offer some challenging thoughts on exchangeability, a central concept in the subjective theory of probability. This note reviews the finite and discrete foundation of de Finetti's analysis of this matter, and identifies its implications for applied statistical research. Currently accepted concepts of Bayesian statistics focus on the representation of exchangeable distributions over infinite sequences of events. Taken seriously, this has led to misplaced efforts, unduly influenced by objectivist statistical constructs associated with the Neyman-Pearson persuasion. My conclusion from these considerations supports the practice of scored sequential forecasting as a sensible alternative to the outdated practice of hypothesis testing.

Key Words: Finite exchangeable inference, exchangeable extendibility, mixture hypergeometric distributions, de Finetti representation theorem, proper scoring rules, hypothesis testing, parameter estimation, Bayesian methods

0 Reminiscence and theme

The Summer School at Asti has provided memorable service in adult education over the years on many fronts. While the focus of the school has always been on applied and computational statistics, course material has been embedded with mathematical and foundational content at an introductory level as well. We have regularly examined the distinction between subjectivist and objectivist views of probability and its relevance to appropriate methods of statistical practice. An introduction to Bayesian statistical methods has regularly been a regular feature of the program. Nonetheless, whatever one feels about the relative merits of various attitudes toward statistical inference, it cannot be denied that the objectivist outlook continues to dominate the general public perception and application of statistical inference today. Even the inroads made by Bayesian inference over the past few decades have been formulated in terms of objectivist concepts.

The common Bayesian characterization of exchangeable inference suffers from its embedment within a statistical framework designed in these terms. Quite distinctly, the substantive foundation of exchangeability is set within a subjectivist tradition originating in the imaginative outlook and the mathematical constructions of Bruno de Finetti. In

A bootstrap test for unit roots

Isabella Procidano and Silio Rigatti-Luchini

Department of Management, University Ca' Foscari of Venice, Italy
Department of Statistical Sciences, University of Padua, Italy

Abstract

This paper presents an external bootstrap test for unit roots in an AR(1) model with heteroschedastic disturbance term. Simulations evidence that the proposed test has a right proportion of rejects also if the variance of the disturbance terms is not constant.

Keywords: Time series; Unit roots; External bootstrap; Dickey-Fuller test.

1 Introduction

In recent years, the literature regarding testing for unit roots has developed exponentially, mainly due to two interrelated factors: the empirical observation that almost all macro-economic time series are affected by stochastic trends, and the importance taken on by the concept of cointegration (Engle and Granger, 1987) in theoretical and applied econometrics. One of the tests most frequently used to verify the presence of unit roots is the Dickey-Fuller (DF) test (Fuller, 1976, 1995; Dickey and Fuller, 1979; Mackinnon, 2010). The hypotheses underlying the applicability of the DF test are the incorrelation and homoskedasticity of the error terms in the model. Violation of the first hypothesis may be treated by using the Augmented Dickey Fuller (ADF) test[1] (Dickey and Fuller, 1981). Instead, violation of the hypothesis of constant variance only recently has it been appropriately treated (Procidano and Rigatti-Luchini, 2002; Pizzi *et al*., 2003; Park, 2003; Cavaliere, 2005; Cavaliere *et al*., 2017), due to the presumed asymptotic robustness of the DF test (Phillips, 1987; Xiao, 2014). In the present paper we complete our previous works using simulations to assess the effects of heteroschedasticity on the effective rejection probabilities of the DF test in time series of finite length, and propose a robust bootstrap test to verify unit roots.

The paper is organized as follows. Section 2 introduces the model, tests, and bootstrap method. The simulation experiment are derived in Section 3. The simulation results are reported in Section 4, and Section 5 concludes the paper.

2. The external bootstrap for testing unit roots

[1] Although the problem arises of determining delays to be taken into account and the consequent loss of observations.

Modern Likelihood Inference for Measures of Process Capability

Erlis Ruli and Laura Ventura

Department of Statistical Sciences, University of Padova, Italy

ruli@stat.unipd.it, ventura@stat.unipd.it

Abstract

Process capability (PC) indices are essential for assessing the capability of manufacturing processes. Point and interval estimation based on classical likelihood procedures about PC indices has been widely examined under various model assumptions. However, it is well-known that first order inference can be inaccurate, in particular when the sample size is small. This situation is common in ECO labelling declaration.

In this paper we illustrate higher-order likelihood-based procedures for parametric inference in small samples, which provide accurate point estimators and confidence intervals for measures of PC. The connection with Bayesian inference based on matching priors is highlighted. Two examples are discussed and an implementation of the proposed methods in the `R` software is provided.

Keywords and phrases. C_p index, higher-order likelihood inference, labelling declaration, matching priors, percentage of non-conforming, quality control, small sample size.

1 Introduction

Process capability (PC) indices, as measures of process performance, have become very popular in assessing the capability of manufacturing processes in practice. They are widely used to determine whether a process is capable of producing items within a specific tolerance; see, among others, Kotz and Johnson (1993), Kotz *et al.* (1993), Kotz and Johnson (2002), Wu *et al.* (2009), and references therein. For Bayesian inference on CP indices see also Shiau *et al.* (1999), Pearn and Wu (2005), Pearn and Kotz (2006), and references therein.

The usual practice in the study of a process performance is to estimate the PC indices from the observed sample and then to judge the capability of the process by these estimates. Commonly used point and interval estimators of PC indices are based

Edizioni ETS
Palazzo Roncioni - Lungarno Mediceo, 16, I-56127 Pisa
info@edizioniets.com - www.edizioniets.com
Finito di stampare nel mese di ottobre 2018